



**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
ВЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
НАУЧНО-ОБРАЗОВАТЕЛЬНЫЙ ЦЕНТР
СУПЕРВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ И СИСТЕМ**

В.С.Князьков, В. В. Афонин, С.А. Федосин

**СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ:
МАТЕМАТИЧЕСКИЕ ОСНОВЫ И МЕТОДЫ АНАЛИЗА
СТОХАСТИЧЕСКИХ ПРОЦЕССОВ И СИСТЕМ**

УЧЕБНОЕ ПОСОБИЕ

Киров 2010

УДК 519.87

ББК 396я73

A946

ISBN 5 – 7103 – 0930 – 3

Системы массового обслуживания: математические основы и методы анализа стохастических процессов и систем. : Учеб. пособие / В.С.Князьков, В. В. Афонин, С.М. ,С.А. Федосин. – Киров, 2010. – 385 с.

В пособии освещаются вопросы теории типовых систем массового обслуживания, которые могут быть описаны процессами размножения и гибели. Изложение материала дается в таком ключе, чтобы можно было подготовиться к экзамену или иной форме теоретической отчетности без дополнительного привлечения литературных источников.

В пособии рассматриваются примеры, решение которых осуществляется с помощью языка программирования системы MATLAB.

Предназначено для студентов и аспирантов различных специальностей.

**УДК 519.87
ББК 396я73**

ISBN 5 – 7103 – 0930 – 3 © В.С.Князьков, В. В. Афонин, С. М.
С. А. Федосин, 2010

ОГЛАВЛЕНИЕ

Введение	4
Глава 1. Анализ потоков требований	11
1.1. Свойство стационарности потока событий	11
1.2. Свойство ординарности потока событий	13
1.3. Свойство отсутствия последействия	14
1.4. Параметр и интенсивность потока требований	15
1.5. Пуассоновский поток требований	18
1.6. Распределение Пуассона	26
1.7. Моделирование пуассоновского потока	43
1.8. Преобразования пуассоновских потоков	44
1.9. Потоки Эрланга	52
Глава 2. Процессы размножения и гибели	70
2.1. Определение и постулаты процесса размножения и гибели	70
2.2. Дифференциальные уравнения процесса размножения и гибели	73
2.3. Процесс чистого размножения	89
2.4. Процесс чистой гибели	94
2.5. Стационарный режим процесса размножения и гибели	99
Глава 3. Однолинейные системы обслуживания	105
3.1. Система массового обслуживания М/М/1. Распределение вероятностей в системе М/М/1	105
3.2. Формула Литтла. Аналоги формулы Литтла	112
3.3. Операционные характеристики системы М/М/1	117
3.4. Система М/М/1 с отказами	128

3.5. Система М/М/1/К: конечный накопитель	132
3.6. Система М/М/1//М с конечным числом источников нагрузки	138
3.7. Система М/М/1 с ограниченным временем ожидания	155
Глава 4. Многолинейные системы обслуживания	180
4.1. Система М/М/∞ с немедленным обслуживанием	181
4.2. Система М/М/т с т обслуживающими приборами	184
4.3. Операционные характеристики системы М/М/т	189
4.4. Система М/М/т с отказами. В-формула Эрланга	200
4.5. Операционные характеристики системы М/М/т с отказами	203
4.6. Пример анализа автоматической телефонной станции как системы массового обслуживания с отказами	206
4.7. Система М/М/т/К с конечным накопителем	212
4.8. Операционные характеристики системы М/М/т/К	216
4.9. Система М/М/т с ограниченным временем ожидания	220
4.10. Система М/М/∞//М с конечным числом источников нагрузки и бесконечным числом обслуживающих приборов	237
4.11. Система М/М/т/К/М с конечным числом источников нагрузки, с т обслуживающими приборами и конечным накопителем	240
4.12. Пример моделирования системы М/М/т/К/М	244
Глава 5. Многофазные системы обслуживания	252
5.1. Двухфазная модель с нулевой вместимостью блока ожидания	252
5.2. Трехфазная модель с нулевой вместимостью блоков	

ожидания	277
5.3. Пример моделирования трехфазной системы массового обслуживания с нулевой вместимостью блоков ожидания	313
Глава 6. Применение смо в задачах анализа производительности многопроцессорных вычислительных систем	322
6.1 Введение	322
6.2. Основы теории стохастических сетей массового обслуживания	332
6.3. Макроанализ производительности микропроцессорных систем	342
6.4. Микроанализ производительности микропроцессорных систем	350
Литература	363

ВВЕДЕНИЕ

В предлагаемом пособии рассматриваются некоторые модели систем массового обслуживания, производится анализ этих моделей в целях получения их операционных характеристик. Для систем массового обслуживания можно привести следующее определение [1]: массового обслуживания система — понятие, которое включает в себя случайный "входящий" поток требований (вызовов, клиентов), нуждающихся в "обслуживании", и механизм (алгоритм), осуществляющий это "обслуживание". Требования, вызовы, клиенты, заявки поступают в некоторые устройства, приборы или каналы обслуживания, где они пребывают в общем случае случайное время. Процесс, характерный для всякой системы массового обслуживания (СМО), показан на рис. В1 [5].

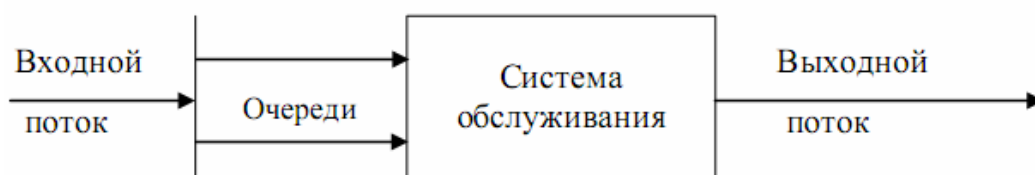


Рис. В1. Структура процесса массового обслуживания

На рис. В1 система обслуживания представляет собой совокупность приборов или каналов обслуживания, которые в общем случае могут иметь сложную структуру. В зависимости от структуры и численности каналов СМО называется однолинейной (одноканальной), многолинейной (многоканальной), однофазной, многофазной. В дальнейшем эти понятия будут раскрыты подробнее.

Для систем массового обслуживания характерным является поступление требований через случайные промежутки времени и случайное время обслуживания этих требований. При этом требования могут создавать очередь перед каналом (каналами) обслуживания. Поступление требований в систему может подчиняться какому-либо вероятностному закону. Случайное время обслуживания также может

быть подчинено вероятностному закону. Создаваемая требованиями очередь может быть ограниченной или неограниченной. Система массового обслуживания может быть с отказами, т. е. когда в ней очередь отсутствует.

Для отражения законов поступления требований, их обслуживания, а также в некотором смысле структуры СМО может быть использована модифицированная символика Кендалла [2, 3]. Основным вид такого символьного обозначения состоит из четырех позиций, разделяемых вертикальными или наклонными линиями: $A/B/m/K$, где

A — распределение моментов поступлений требований на обслуживание;

B — распределение времени обслуживания требований;

m — число параллельно функционирующих узлов (каналов, приборов) обслуживания;

K — максимальное число допускаемых в систему требований, т. е. число требований в очереди плюс число требований, принятых на обслуживание.

Кроме указанных обозначений могут иметь место и такие, которые связаны с емкостью источника нагрузки, генерирующего требования на обслуживание, и с дисциплиной очереди.

Поля A и B в модели СМО могут принимать идентичные значения из следующего набора стандартных обозначений [3]:

M — экспоненциальное (показательное) распределение. Для входящего потока промежутки времени между последовательными поступлениями требований распределены экспоненциально, а поступление самих требований в систему подчиняется закону Пуассона;

D — детерминированный интервал времени между моментами последовательных поступлений в систему требований на обслуживание или детерминированная продолжительность обслуживания;

E_r — распределение Эрланга порядка r интервалов времени между моментами последовательных поступлений требований в обслуживающую систему или продолжительностей обслуживания;

GI — распределение произвольного вида моментов поступления в систему требований на обслуживание (или интервалов времени между последовательными поступлениями требований);

G — распределение произвольного вида моментов выбытия из системы обслуженных требований (или продолжительностей обслуживания).

Если СМО допускает очередь, то говорят, что эта система с накопителем. Например, $M/M/1/K$ — это однолинейная (символ 1) система обслуживания с конечным накопителем (символ K), с пуассоновским входящим потоком (первый символ M) и экспоненциальным каналом обслуживания (второй символ M). Отметим, что символ K означает допустимое число требований, которые могут находиться в системе: на обслуживании (m штук) плюс в очереди ($K-m$). Система с m обслуживающими каналами (приборами, узлами) и с конечным накопителем может быть изображена так, как показано на рис. В2, где H — накопитель.

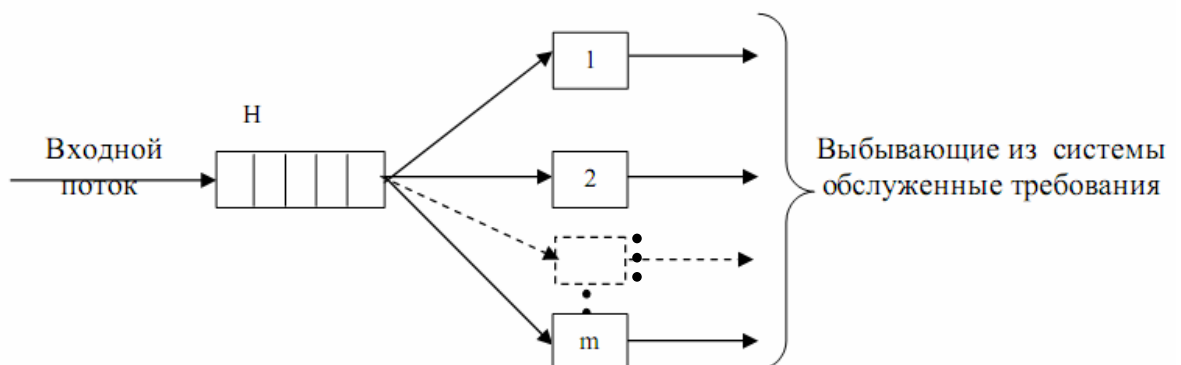


Рис. В2. Схема системы с накопителем и m каналами обслуживания

Систему, приведенную на рис. В2, называют также многолинейной системой массового обслуживания.

СМО может представляться в виде многофазной модели (рис. В3), в которой требование последовательно обслуживается на всех фазах. Перед каждой фазой (1, 2, . . . , N) может допускаться неограниченная очередь [2].

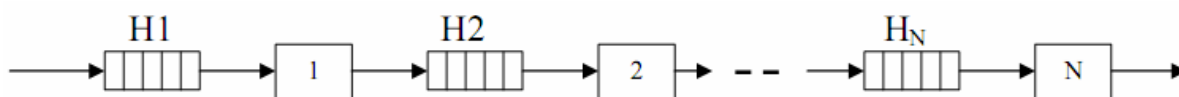


Рис. В3. Схема многофазной СМО

На рис. В4 показана многофазно-многолинейная модель, представляющая собой последовательно расположенные группы (фазы) параллельно работающих приборов [2].

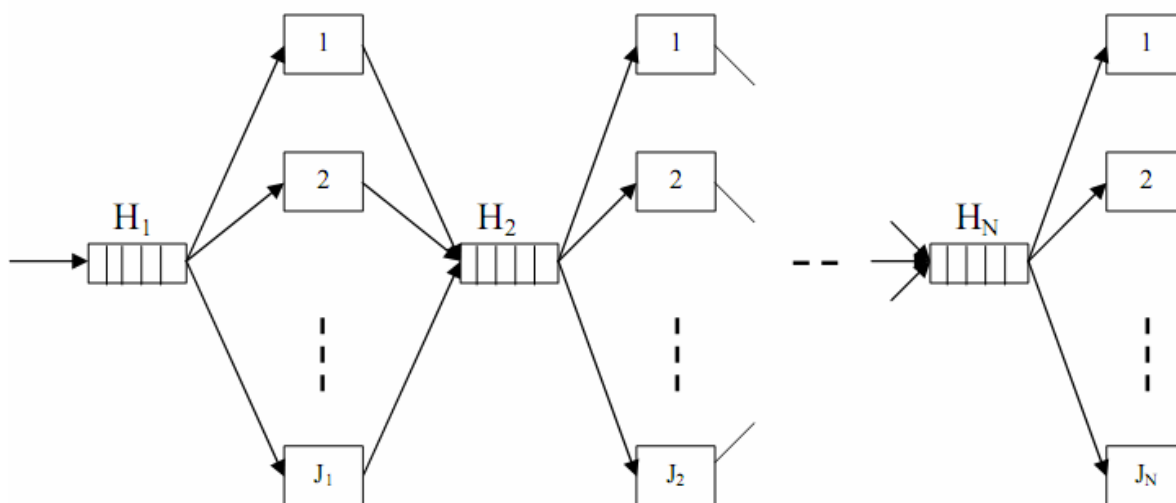


Рис. В4. Схема многофазно-многолинейной СМО

Числа J_1, J_2, \dots, J_N на рис. В4 означают количество приборов на фазах 1, 2, . . . , N соответственно. Поток требований обозначен стрелками. Очереди H_1, H_2, \dots, H_N изображены через накопители.

В символьные обозначения систем массового обслуживания может входить дополнительное поле для определения приоритета обслуживания. Например, для приоритетных систем в пятом поле фиксируется символ f_i^j , $i = 0, 1, 2$; $j = 0, 1$. Если $i = 0$, то

осуществляется обслуживание без приоритета, при $i=1$ в системе имеется относительный приоритет; $i=2$ означает наличие абсолютного приоритета; значение $j=0$ указывает, что требование, заставшее все места занятыми, теряется; $j=1$ — вновь прибывшее требование вытесняет требование с низшим приоритетом [5]. Приоритетное обслуживание применяется в тех случаях, когда в систему поступают разнородные требования. Приведем определение однородных и неоднородных требований. Если требования, поступающие в систему массового обслуживания, различаются только моментами времени поступления в систему, то такие требования называются *однородными*. Если требования имеют дополнительные признаки, по которым формируются приоритетные свойства, то такие требования будут называться *неоднородными*.

Для неоднородных требований возможны следующие дисциплины обслуживания [6].

Относительный приоритет. При завершении обслуживания очередного требования из очереди на обслуживание выбирается требование с наивысшим приоритетом. Относительный приоритет предусматривает, что поступление требования с более высоким приоритетом не прерывает обслуживания менее приоритетного требования.

Абсолютный приоритет. При поступлении в СМО требования более высокого приоритета, чем требование, обслуживаемое на приборе, происходит прерывание обслуживания, начинает обслуживаться поступившее требование, а прерванное либо теряется (что соответствует разрушающему приоритету [5, с.54]), либо возвращается в очередь с последующим дообслуживанием (что соответствует абсолютному приоритету с запоминанием [5, с.54]) или обслуживанием заново (что

соответствует абсолютному приоритету с повторением [5, с.54]). Дисциплины с абсолютными приоритетами могут использоваться только для систем с одним устройством обслуживания.

Чередующийся приоритет. Он предусматривает закрепление за требованиями того типа, который находится на обслуживание наивысшего приоритета. После обслуживания всех имеющихся в СМО требований этого типа из очереди на обслуживание выбирается следующее требование, как правило, в соответствии с относительным приоритетом.

Для систем массового обслуживания с ожиданием и однородными требованиями принято рассматривать следующие схемы определения порядка обслуживания требований [6]: прямой, инверсионный, случайный.

Прямой порядок обслуживания, или обслуживание в порядке поступления в СМО. Для такой дисциплины могут применяться обозначения: FIFO (First In First Out — первым пришел, первым вышел) или FCFS (First Come First Serve — первый поступивший первым обслуживается).

Инверсионный (обратный, стековый) порядок обслуживания. Он характеризуется выбором на обслуживание требования, поступившего в СМО последним. Для такой дисциплины могут применяться обозначения: LIFO (Last In First Out — последним пришел, первым вышел) или LCFS (Last Come First Serve — последний поступивший первым обслуживается).

Случайный выбор требований для обслуживания. При случайном порядке обслуживания с равной вероятностью может быть выбрано на обслуживание любое из имеющихся в очереди требований. Подразумевается равномерный закон распределения вероятностей. Для

такой дисциплины может применяться обозначение RANDOM (случайно).

Дисциплины постановки в очередь и выбора из нее определяют порядок постановки требований в очередь, если заняты устройства обслуживания, и порядок выбора из очереди, если освобождается обслуживающее устройство.

ГЛАВА 1. АНАЛИЗ ПОТОКОВ ТРЕБОВАНИЙ

Для каждой системы массового обслуживания неотъемлемой составляющей частью является наличие потока требований: входного потока требований, обслуженного потока, потока требований, получивших отказ в обслуживании, выходного потока требований. В соответствии со структурой процесса массового обслуживания (рис. В1) естественно начать рассмотрение с входного потока требований, которые поступают в СМО в соответствии с каким-либо законом. В общем случае входной поток требований является вероятностным потоком, и поэтому все его характеристики также будут вероятностными. К числу важнейших характеристик вероятностного потока относятся:

- вероятность поступления того или иного числа требований на заданном отрезке времени;
- среднее число требований, поступающих за данное время;
- вероятностное распределение длин временных интервалов между соседними требованиями;
- средний интервал времени между двумя соседними требованиями;
- интенсивность поступления требований в систему [7].

Считая требования однородными, входящий поток будет определяться как некоторое множество моментов времени [5]. С точки зрения поступления требований в систему можно сказать, что в этой системе происходят события. Тогда множество моментов, когда происходят события, называется потоком однородных событий [5]. В реальных системах массового обслуживания события происходят в случайные моменты времени, поэтому следует говорить о случайном

потоке событий. Отсюда напрашивается следующее определение: случайный поток однородных событий — это поток однородных событий, зависящий от случая. Математически строгое определение приведено в [5, с.13-14].

Среди основных свойств потоков событий важное место занимают свойства стационарности, ординарности, отсутствия последействия.

1.1. Свойство стационарности потока событий

Введем следующие обозначения [7]:

$P_k(t_0, t)$ — вероятность того, что на отрезке времени с началом в точке t_0 и длиной t (т. е. на отрезке $[t_0; t_0 + t]$, поступило ровно k требований из данного потока;

$P_{\geq k}(t_0, t)$ — вероятность того, что на данном отрезке времени поступит не менее k требований;

$P_{\leq k}(t_0, t)$ — вероятность того, что на данном отрезке времени поступит не более k требований;

$P_0(t_0, t)$ — вероятность отсутствия требований на данном отрезке времени.

Поток требований называется *стационарным*, если для любого числа требований k и двух отрезков времени одинаковой длины $[t_0, t_0 + t]$, $[t_1, t_1 + t]$ выполнено равенство

$$P_k(t_0, t) = P_k(t_1, t). \quad (1.1)$$

Для стационарного потока требований условие (1.1) означает, что вероятность поступления k требований в систему зависит от длины временного интервала и не зависит от того, где именно на оси времени этот интервал расположен [8].

Для стационарных потоков вероятность поступления точно k требований на отрезке времени t обозначается как $P_k(t)$.

1.2. Свойство ординарности потока событий

Поток событий называется *ординарным*, если вероятность попадания на элементарный участок Δt двух или более событий пренебрежимо мала по сравнению с вероятностью попадания одного события.

В соответствии с введенными обозначениями вероятностей приведем еще одно определение [7]: поток называется ординарным, если он удовлетворяет условию: какова бы ни была точка t_0 ,

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t_0, t)}{t} = 0. \quad (1.2)$$

Выражение (1.2) означает, что вероятность поступления двух или большего числа требований на отрезке $[t_0, t_0 + t]$ стремится к нулю (при $t \rightarrow 0$) и при этом существенно быстрее, чем сама длина этого отрезка t .

Предельное соотношение (1.2) эквивалентно условию

$$P_{\geq 2}(t_0, t) = o(t), \quad (1.3)$$

где $o(t)$ — бесконечно малая величина более высокого порядка, чем t .

В случае стационарного потока условие ординарности вида (1.2) или (1.3) приобретает соответственно вид

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} = 0, \quad (1.4)$$

$$P_{\geq 2}(t) = o(t), \quad (1.5)$$

Для любого t имеем следующую полную группу событий:

$$P_0(t_0, t) + P_1(t_0, t) + P_{\geq 2}(t_0, t) = 1. \quad (1.6)$$

В случае ординарного потока при $t \rightarrow 0$ будем иметь приближенное равенство

$$P_0(t_0, t) + P_1(t_0, t) \approx 1. \quad (1.7)$$

В случае стационарного потока (1.7) перепишется в виде

$$P_0(t) + P_1(t) \approx 1.$$

1.3. Свойство отсутствия последействия

Поток событий называется потоком без последействия, если для любых неперекрывающихся участков времени число событий, попадающих на один из них, не зависит от числа событий, попадающих на другие [9].

Приведем более строгое определение. *Поток требований будет являться потоком без последействия*, если условные вероятности поступления k требований на произвольном отрезке времени $[t_0, t_0 + t]$, вычисленные при различных предположениях о распределении моментов поступления требований до t_0 , совпадают с безусловной вероятностью $P_k(t_0, t)$.

Отсутствие последействия состоит в том, что вероятность поступления k требований в течение произвольного промежутка времени не зависит от того, сколько требований и как поступали до этого промежутка.

Наличие последействия означает, что поступление требований на тех или иных промежутках времени влияет на вероятности возникновения дальнейших требований на других промежутках времени [7].

Отметим, что случайный процесс (в данном случае процесс поступления требований) без последействия называется *Марковским*

процессом (по имени русского ученого А. А. Маркова).

1.4. Параметр и интенсивность потока требований

Параметр и интенсивность являются характеристиками потока требований.

1.4.1. Параметр и интенсивность стационарного потока

Параметром стационарного потока называется предел

$$\lim_{t \rightarrow 0} \frac{1 - P_0(t)}{t} = \lim_{t \rightarrow 0} \frac{P_{\geq 1}(t)}{t} = a. \quad (1.8)$$

Из предельного соотношения (1.8) можно записать

$$1 - P_0(t) = P_{\geq 1}(t) = at + o(t), \quad (1.9)$$

где $P_0(t)$ — вероятность того, что на отрезке времени t требований не будет, $P_{\geq 1}(t)$ — вероятность того, что на отрезке времени t будет по крайней мере одно требование (и больше одного), $o(t)$ — бесконечно малая величина более высокого порядка, чем t .

Параметр потока показывает скорость сходимости к нулю вероятности поступления хотя бы одного требования на отрезке времени t при стремлении к нулю длины этого отрезка [7]. Если от обеих частей равенства (1.9) взять предел при $t \rightarrow 0$, то получим

$$\lim_{t \rightarrow 0} P_{\geq 1}(t) = \lim_{t \rightarrow 0} at + \lim_{t \rightarrow 0} o(t),$$

откуда

$$P_{\geq 1}(0) = 0. \quad (1.10)$$

Соотношение (1.10) означает, что вероятность поступления хотя бы одного требования в точке равна нулю.

Беря пределы в (1.9) относительно $P_0(t)$, получим

$$P_0(0) = 1, \quad (1.11)$$

т. е. вероятность отсутствия требований в точке равна 1. Следовательно, условие полной группы событий принимает вид:

$$P_0(0) + P_{\geq 1}(0) = 1. \quad (1.12)$$

Интенсивностью стационарного потока называется среднее число требований, поступающих из потока за единицу времени [7]. Приведенное определение соответствует вычислению математического ожидания случайной дискретной величины. Обозначим интенсивность потока через λ , тогда

$$\lambda = \sum_{k=0}^{\infty} k P_k(1), \quad (1.13)$$

где $P_k(1)$ — вероятность поступления точно k требований на отрезке времени длиной в одну относительную единицу.

Для отрезка времени произвольной величины интенсивность стационарного потока будет вычисляться в виде:

$$\lambda = \frac{\sum_{k=0}^{\infty} k P_k(t)}{t}.$$

В стационарном случае параметр и интенсивность потока являются постоянными не отрицательными величинами.

1.4.2. Параметр и интенсивность нестационарного потока

Для нестационарного потока требований значение параметра может меняться во времени [7]. *Мгновенным значением параметра* в момент времени t_0 называется предел

$$\lim_{t \rightarrow 0} \frac{1 - P_0(t_0, t)}{t} = \lim_{t \rightarrow 0} \frac{P_{\geq 1}(t_0, t)}{t} = a(t_0). \quad (1.14)$$

Для нестационарного потока вводится понятие мгновенной интенсивности. Мгновенным значением интенсивности в момент t_0 называется предел

$$\lim_{t \rightarrow 0} \frac{\sum_{k=0}^{\infty} kP_k(t_0, t)}{t} = \lambda(t_0). \quad (1.15)$$

Числитель (1.15) — это среднее число требований, поступающих на отрезке времени $[t_0, t_0 + t]$.

Мгновенные значения параметра и интенсивности связаны соотношением [7]:

$$\lambda(t_0) \geq a(t_0). \quad (1.16)$$

Покажем справедливость соотношения (1.16). В числителе (1.15) нулевое слагаемое можно не учитывать, поэтому получаем очевидное соотношение:

$$\sum_{k=0}^{\infty} kP_k(t_0, t) = \sum_{k=1}^{\infty} kP_k(t_0, t). \quad (1.17)$$

По теореме о сложении несовместных вероятностей можем переписать числитель (1.14) в виде

$$P_{\geq 1}(t_0, t) = \sum_{k=1}^{\infty} P_k(t_0, t). \quad (1.18)$$

Предельные соотношения (1.14) и (1.15) различаются только числителями, которые определены через (1.17) и (1.18). Из сравнения (1.17) и (1.18) следует нестрогое неравенство

$$\sum_{k=1}^{\infty} kP_k(t_0, t) \geq \sum_{k=1}^{\infty} P_k(t_0, t), \quad (1.19)$$

поскольку левая сумма (1.19) это

$$1P_1(t_0, t) + 2P_2(t_0, t) + 3P_3(t_0, t) + \dots,$$

а правая сумма (1.19) есть

$$P_1(t_0, t) + P_2(t_0, t) + P_3(t_0, t) + \dots$$

Далее разделим обе части неравенства (1.19) на t , возьмем предел при $t \rightarrow 0$ и получим неравенство (1.16), т. е. $\lambda(t_0) \geq a(t_0)$ — интенсивность потока не меньше его параметра. Для стационарных потоков $\lambda \geq a$.

1.5. Пуассоновский поток требований

Поток требований называется *пуассоновским*, если он обладает свойствами ординарности, стационарности и отсутствия последействия [7]. В общем случае поток будет пуассоновским, если он обладает только свойствами ординарности и отсутствия последействия. Стационарный пуассоновский поток называется *простейшим* [19]. В дальнейшем под пуассоновским потоком будем понимать простейший поток.

Простейший поток имеет большое значение в приложениях по той причине, что при наложении (суперпозиции) достаточно большого числа n независимых, стационарных и ординарных потоков (сравнимых между собой по интенсивностям $\lambda_i, (i = 1, 2, \dots, n)$) получается поток, близкий к простейшему с интенсивностью λ , равной сумме интенсивностей входящих потоков [10].

Поставим и решим задачу определения вероятности того, что в течение промежутка времени $t + \Delta t$ поступит точно k требований из пуассоновского потока — ординарного, стационарного, с отсутствием последействия. Эту вероятность обозначим как $P_k(t + \Delta t)$. Событие в появлении точно k требований на отрезке времени $t + \Delta t$ может произойти $k+1$ различными способами:

1-й способ. За время t поступят все k требований, за Δt — ни одного.

2-й способ. За время t поступят $k-1$ требований, за Δt — одно.

3-й способ. За время t поступят $k-2$ требований, за Δt — два.

.....

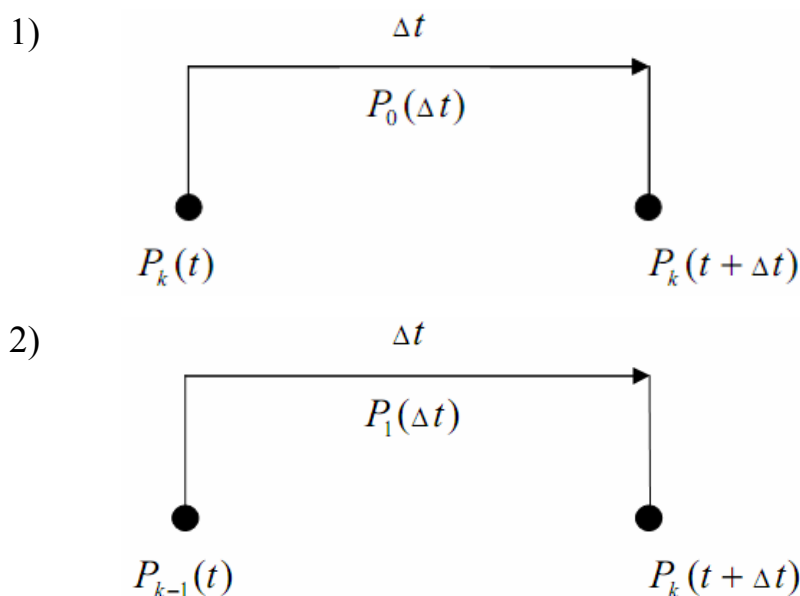
k -й способ. За время t поступит одно требование, за Δt — $k-1$ требований.

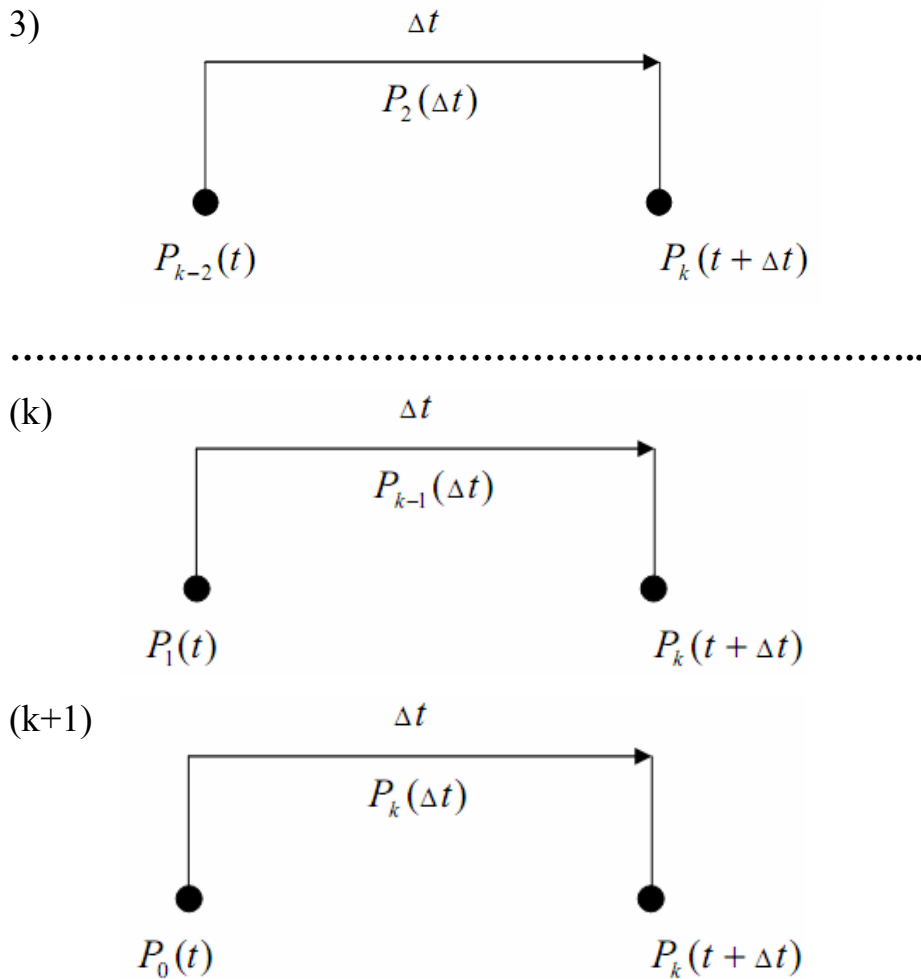
$(k+1)$ -й способ. За время t не поступит ни одного требования, за Δt — все k требований.

Из определения параметра стационарного потока, можем записать

$$1 - P_0(t) = at + o(t), \Rightarrow P_0(t) = 1 - at + o(t),$$

где бесконечно малая величина $o(t)$ взята со знаком плюс, так как прибавление или вычитание бесконечно малой величины не изменяет конечное число с точностью до бесконечно малых более высокого порядка, чем t . Приведенные способы поступления требований можно проиллюстрировать следующими диаграммами:





Поступления требований на заданном отрезке времени — это независимые события. Такое допущение оправдано для потока однородных событий, следующих одно за другим в какие-то случайные моменты времени.

Напомним некоторые сведения из теории вероятностей. *События A и B называются несовместными*, если наступление одного из них исключает наступление другого [11, с.19]. *Два события называются независимыми*, если появление одного из них не меняет вероятности появления другого [10]. *События A_1, A_2, \dots, A_n образуют полную группу событий*, если они попарно несовместны и вместе образуют достоверное событие [11]. *Правило сложения вероятностей*: вероятность суммы несовместных событий равна сумме вероятностей этих событий, т. е. $P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$ [10, с.35]. *Правило*

умножения вероятностей для независимых событий: вероятность произведения двух или нескольких независимых событий A, B, \dots, N равна произведению вероятностей этих событий [10, с.41-42], т. е. $P(AB\dots N) = P(A)P(B)\dots P(N)$. Следствием правил (теорем) сложения и умножения вероятностей является *теорема о формуле полной вероятностей*: если события A_1, A_2, \dots, A_n образуют полную группу событий, то вероятность события B может быть представлена как сумма произведений безусловных вероятностей событий полной группы на условные вероятности события B :

$$P(B) = \sum_{i=1}^n P(A_i)P(B / A_i), \quad (1.20)$$

где $P(B / A_i)$ — условная вероятность, т. е. вероятность события B , найденная при условии, что событие A_i произошло.

Вернемся к поставленной задаче о нахождении вероятности появления k требований на отрезке времени $t + \Delta t$. В силу свойства отсутствия последействия рассматриваемого потока (простейшего) вытекает, что вероятность попадания любого числа требований на один из временных участков не зависит от того, сколько их попало на другие участки. Это в свою очередь приводит к тому, что условная вероятность попадания требований на какой-либо временной отрезок будет равна безусловной вероятности. Тогда формула полной вероятности (1.20) для рассматриваемой задачи будет иметь вид

$$\begin{aligned} P_k(t + \Delta t) &= \sum_{m=0}^k P_m(t)P_{k-m}(\Delta t) = \\ &= P_0(t)P_k(\Delta t) + P_1(t)P_{k-1}(\Delta t) + \dots + P_{k-1}(t)P_1(\Delta t) + P_k(t)P_0(\Delta t). \end{aligned} \quad (1.21)$$

Выражение (1.21) реализует перечисленные выше условия появления k требований на отрезке времени $t + \Delta t$, т. е. $[t; t + \Delta t]$.

Поскольку рассматривается ординарный поток, а также стационарный, то

в соответствии с условием ординарности (1.5), можем для малого отрезка времени Δt условие ординарности потока записать в виде:

$$P_{\geq 2}(\Delta t) = o(\Delta t), \quad (1.22)$$

где $o(\Delta t)$ — бесконечно малая величина более высокого порядка, чем Δt .

Принимая во внимание условие ординарности (1.22), в выражении (1.21) все слагаемые, содержащие $P_{\geq 2}(\Delta t)$, обратятся в малые величины $o(\Delta t)$, которые объединим в одну малую величину, поскольку сумма бесконечно малых величин есть бесконечно малая величина. Перепишем (1.21) с учетом (1.22):

$$P_k(t + \Delta t) = \sum_{m=0}^k P_m(t) P_{k-m}(\Delta t) = P_{k-1}(t) P_1(\Delta t) + P_k(t) P_0(\Delta t) + o(\Delta t). \quad (1.23)$$

Введем в рассмотрение параметр потока в соответствии с выражением (1.9), заменяя величину t на Δt :

$$1 - P_0(\Delta t) = a\Delta t + o(\Delta t), \quad \Rightarrow \quad P_0(\Delta t) = 1 - a\Delta t + o(\Delta t). \quad (1.24)$$

Далее обратимся к полной группе событий для вероятностей, зависящих от отрезка времени Δt :

$$P_0(\Delta t) + P_1(\Delta t) + P_{\geq 2}(\Delta t) = 1. \quad (1.25)$$

Учитывая (1.22) и (1.24), будем иметь:

$$P_1(\Delta t) = 1 - [1 - a\Delta t + o(\Delta t)] + o(\Delta t), \quad \Rightarrow \quad P_1(\Delta t) = a\Delta t + o(\Delta t), \quad (1.26)$$

где $o(\Delta t)$ — суммарная бесконечно малая величина.

Вернемся к выражению (1.23), куда подставим (1.24) и (1.26):

$$P_k(t + \Delta t) = P_{k-1}(t)[a\Delta t + o(\Delta t)] + P_k(t)[1 - a\Delta t + o(\Delta t)] + o(\Delta t). \quad (1.27)$$

Раскроем выражение (1.27), учитывая, что слагаемые с сомножителями $o(\Delta t)$ становятся бесконечно малыми величинами. При этом получаемые малые величины просуммируем и обозначим как $o(\Delta t)$:

$$P_k(t + \Delta t) = P_{k-1}(t)a\Delta t + P_k(t) - P_k(t)a\Delta t + o(\Delta t). \quad (1.28)$$

В выражении (1.28) вероятность $P_k(t)$ перенесем в левую часть, поделим

обе части выражения на Δt и найдем от обеих частей предел при $\Delta t \rightarrow 0$:

$$\lim_{\Delta t \rightarrow 0} \frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} [P_{k-1}(t)a] - \lim_{\Delta t \rightarrow 0} [P_k(t)a] + \lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t}. \quad (1.29)$$

В выражении (1.29) предел отношения бесконечно малой величины $o(\Delta t)$ к Δt есть ноль. Левая часть выражения (1.29) — это производная, т. е.

$$\lim_{\Delta t \rightarrow 0} \frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \frac{dP_k(t)}{dt}. \quad (1.30)$$

В (1.29) пределы от величин, не содержащих Δt , будут равны этим величинам. В итоге получаем следующую систему обыкновенных дифференциальных уравнений с постоянными коэффициентами:

$$\frac{dP_k(t)}{dt} = aP_{k-1}(t) - aP_k(t), \quad k \geq 1. \quad (1.31)$$

При $k = 0$ из (1.31) формально получаем

$$\frac{dP_0(t)}{dt} = aP_{-1}(t) - aP_0(t), \quad k = 0. \quad (1.32)$$

В силу того, что анализируется поток однородных событий, представляющих собой множество натуральных чисел, то вероятность $P_{-1}(t) = 0$, как вероятность несуществующего события. Таким образом, приходим к следующему бесконечному множеству систем дифференциальных уравнений, описывающих пуассоновский поток:

$$\left. \begin{aligned} \frac{dP_0(t)}{dt} &= -aP_0(t), \\ \frac{dP_k(t)}{dt} &= aP_{k-1}(t) - aP_k(t), \\ k &\geq 1. \end{aligned} \right\} \quad (1.33)$$

Для решения системы (1.33) следует задать начальные условия. Можно, например, использовать естественные начальные условия, т. е.

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k \geq 1. \quad (1.34)$$

Естественные граничные условия (1.34) могут быть получены следующим образом. Решим в (1.33) дифференциальное уравнение относительно $P_0(t)$:

$$\frac{dP_0(t)}{dt} = -aP_0(t), \quad dP_0(t) = -aP_0(t)dt, \quad \frac{dP_0(t)}{P_0(t)} = -adt, \\ \ln(P_0(t)) = -at + \ln C, \quad (1.35)$$

где $\ln C$ — произвольная постоянная интегрирования, взятая как логарифм натуральный от постоянной C .

Потенцируя выражение (1.35), получим

$$P_0(t) = Ce^{-at}. \quad (1.36)$$

Для нахождения постоянной интегрирования в (1.36) необходимо знать вероятность отсутствия требований (событий) на момент времени $t = 0$. На этом этапе предположим, что $P_0(0) = 1$. Тогда выражение (1.36) принимает вид

$$t = 0: P_0(0) = 1, \Rightarrow C = 1, \Rightarrow P_0(t) = e^{-at}. \quad (1.37)$$

Теперь используем условие полной группы событий на момент времени $t = 0$:

$$\sum_{k=0}^{\infty} P_k(0) = 1, \quad P_0(0) + \sum_{k=1}^{\infty} P_k(0) = 1, \quad \{P_0(0) = 1\}1 + \sum_{k=1}^{\infty} P_k(0) = 1, \quad \sum_{k=1}^{\infty} P_k(0) = 0. \quad (1.38)$$

Учитывая, что вероятности — это положительные величины, из (1.38) получаем естественные граничные условия (1.34).

Теперь покажем, что в пуассоновском потоке $a = \lambda$, т. е. параметр и интенсивность равны между собой.

Для стационарного потока выпишем соотношение для определения интенсивности потока:

$$\lambda = \frac{\sum_{k=0}^{\infty} kP_k(t)}{t}. \quad (1.39)$$

Запишем (1.39) в виде

$$\lambda t = \sum_{k=0}^{\infty} kP_k(t) = 0P_0(t) + P_1(t) + 2P_2(t) + 3P_3(t) \dots \quad (1.40)$$

Используем выражение (1.9) для определения параметра стационарного потока при $t \rightarrow 0$:

$$1 - P_0(t) = at + o(t), \Rightarrow at = 1 - P_0(t) + o(t), \left\{ P_0(t) = 1 - \sum_{k=1}^{\infty} P_k(t) \right\} \Rightarrow$$

$$at = \sum_{k=1}^{\infty} P_k(t) + o(t), \Rightarrow at = P_1(t) + P_2(t) + P_3(t) + \dots + o(t).$$

Поскольку мы рассматриваем ординарный поток, то все вероятности, начиная с номера 2 и выше, обращаются в бесконечно малую величину более высокого порядка малости, чем t . Поэтому можем записать

$$at = P_1(t) + o(t). \quad (1.41)$$

Если теперь выражение (1.40) также рассматривать при $t \rightarrow 0$, то также все вероятности, начиная с номера 2 и выше, обращаются в бесконечно малую величину более высокого порядка малости, чем t в силу ординарности потока, для которого рассматривается его интенсивность. Тогда получаем:

$$\lambda t = P_1(t) + o(t). \quad (1.42)$$

Из сравнения (1.41) и (1.42) заключаем, что в пуассоновском потоке его параметр равен его интенсивности, т. е.

$$a = \lambda. \quad (1.43)$$

Учитывая равенство (1.43), заменим в системе дифференциальных уравнений параметр потока на его интенсивность:

$$\left. \begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda P_0(t), \\ \frac{dP_k(t)}{dt} &= \lambda P_{k-1}(t) - \lambda P_k(t), \\ k &\geq 1. \end{aligned} \right\} \quad (1.44)$$

Граничные условия для решения (1.44) по-прежнему можно считать естественными, т. е.

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k \geq 1. \quad (1.45)$$

В дальнейшем, при рассмотрении пуассоновского (простейшего) потока будем считать, что в качестве его параметра принимается его интенсивность. Следует иметь в виде, что для пуассоновского потока должно выполняться нормировочное условие вида

$$\sum_{k=0}^{\infty} P_k(t) = 1, \quad (1.46)$$

где $P_k(t)$ может рассматриваться для любого момента времени.

1.6. Распределение Пуассона

1.6.1. Решение дифференциальных уравнений пуассоновского потока

Дифференциальные уравнения (1.44), описывающие поток Пуассона являются обыкновенными дифференциальными уравнениями с постоянными коэффициентами. Рассмотрим решения системы дифференциальных уравнений (1.44) с начальными условиями (1.45). Решение для $P_0(t)$ нам известно, оно определяется выражением

$$P_0(t) = e^{-\lambda t}. \quad (1.47)$$

Обратимся к следующему уравнению системы (1.44) при $k = 1$:

$$\frac{dP_1(t)}{dt} = \lambda P_0(t) - \lambda P_1(t). \quad (1.48)$$

Используя решение (1.47), перепишем (1.48) в виде

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t) + \lambda e^{-\lambda t}. \quad (1.49)$$

Уравнение (1.49) — это неоднородное дифференциальное линейное уравнение. Его решение определяется как сумма решений общего однородного уравнения и частного неоднородного уравнения. Для нахождения общего решения неоднородного уравнения применим метод вариации произвольных постоянных — метод Лагранжа [12]. Сначала находим общее решение соответствующего однородного уравнения (1.49):

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t); \quad \frac{dP_1(t)}{P_1(t)} = -\lambda dt; \quad \ln P_1(t) = -\lambda t + \ln C_1.$$

После потенцирования логарифмического уравнения получаем

$$P_1(t) = C_1 e^{-\lambda t},$$

где C_1 — постоянная интегрирования.

Полагаем, что постоянная интегрирования зависит от времени, т. е. $C_1(t)$.

Тогда запишем решение в виде

$$P_1(t) = C_1(t) e^{-\lambda t}. \quad (1.50)$$

Продифференцируем уравнение (1.50) по времени t :

$$\frac{dP_1(t)}{dt} = \frac{dC_1(t)}{dt} e^{-\lambda t} - C_1(t) \lambda e^{-\lambda t}. \quad (1.51)$$

Сопоставляя (1.49) и (1.51), будем иметь

$$\frac{dC_1(t)}{dt} e^{-\lambda t} - C_1(t) \lambda e^{-\lambda t} = -\lambda C_1(t) e^{-\lambda t} + \lambda e^{-\lambda t}.$$

После приведения подобных членов и сокращения на экспоненту получим

$$\frac{dC_1(t)}{dt} = \lambda; \quad dC_1(t) = \lambda dt; \quad C_1(t) = \lambda t + c_1,$$

где c_1 — постоянная интегрирования, подлежащая определению из начальных условий.

Полученное решение относительно $C_1(t)$ подставим в (1.50):

$$P_1(t) = (\lambda t + c_1)e^{-\lambda t} = \lambda te^{-\lambda t} + c_1 e^{-\lambda t}.$$

Для определения постоянной интегрирования обратимся к начальным условиям (1.45) относительно $P_1(t)$:

$$t = 0; \quad P_1(0) = 0.$$

Откуда следует, что $c_1 = 0$.

Решение для $P_1(t)$ будет иметь вид

$$P_1(t) = \lambda te^{-\lambda t}. \quad (1.52)$$

Обратимся к следующему уравнению системы (1.45) при $k = 2$:

$$\frac{dP_2(t)}{dt} = -\lambda P_2(t) + \lambda P_1(t).$$

Используя решение (1.52), будем иметь

$$\frac{dP_2(t)}{dt} = -\lambda P_2(t) + \lambda^2 te^{-\lambda t}. \quad (1.53)$$

Находим общее решение соответствующего однородного уравнения (1.53):

$$\frac{dP_2(t)}{dt} = -\lambda P_2(t); \quad \frac{dP_2(t)}{P_2(t)} = -\lambda dt; \quad \ln P_2(t) = -\lambda t + \ln C_2.$$

После потенцирования логарифмического уравнения получаем

$$P_2(t) = C_2 e^{-\lambda t},$$

где C_2 — постоянная интегрирования.

Полагаем, что постоянная интегрирования зависит от времени, т. е. $C_2(t)$.

Тогда запишем решение в виде

$$P_2(t) = C_2(t)e^{-\lambda t}. \quad (1.54)$$

Продифференцируем уравнение (1.54) по времени t :

$$\frac{dP_2(t)}{dt} = \frac{dC_2(t)}{dt} e^{-\lambda t} - C_2(t) \lambda e^{-\lambda t}. \quad (1.55)$$

Сопоставляя (1.53) и (1.55), будем иметь

$$\frac{dC_2(t)}{dt} e^{-\lambda t} - C_2(t) \lambda e^{-\lambda t} = -\lambda C_2(t) e^{-\lambda t} + \lambda^2 t e^{-\lambda t}.$$

После приведения подобных членов и сокращения на экспоненту получим:

$$\frac{dC_2(t)}{dt} = \lambda^2 t; \quad dC_2(t) = \lambda^2 t dt; \quad C_2(t) = \frac{(\lambda t)^2}{2} + c_2,$$

где c_2 — постоянная интегрирования, подлежащая определению из начальных условий.

Полученное решение относительно $C_2(t)$ подставим в (1.54):

$$P_2(t) = \left(\frac{(\lambda t)^2}{2} + c_2 \right) e^{-\lambda t} = \frac{(\lambda t)^2}{2} e^{-\lambda t} + c_2 e^{-\lambda t}.$$

Для определения постоянной интегрирования обратимся к начальным условиям (1.45) относительно $P_2(t)$:

$$t = 0; \quad P_2(0) = 0.$$

Откуда следует, что $c_2 = 0$.

Решение для $P_2(t)$ будет иметь вид

$$P_2(t) = \frac{(\lambda t)^2}{2} e^{-\lambda t}. \quad (1.56)$$

Решим еще одно уравнение системы (1.44) при $k = 3$. Исходное уравнение имеет вид

$$\frac{dP_3(t)}{dt} = -\lambda P_3(t) + \lambda P_2(t).$$

С учетом известного решения (1.56) для $P_2(t)$ имеем следующее неоднородное уравнение:

$$\frac{dP_3(t)}{dt} = -\lambda P_3(t) + \lambda \frac{(\lambda t)^2}{2} e^{-\lambda t}. \quad (1.57)$$

Находим общее решение соответствующего однородного уравнения (1.57):

$$\frac{dP_3(t)}{dt} = -\lambda P_3(t); \quad \frac{dP_3(t)}{P_3(t)} = -\lambda dt; \quad \ln P_3(t) = -\lambda t + \ln C_3.$$

После потенцирования логарифмического уравнения получаем:

$$P_3(t) = C_3 e^{-\lambda t},$$

где C_3 — постоянная интегрирования.

Полагаем, что постоянная интегрирования зависит от времени, т. е. $C_3(t)$.

Тогда запишем решение в виде

$$P_3(t) = C_3(t) e^{-\lambda t}. \quad (1.58)$$

Продифференцируем уравнение (1.58) по времени t :

$$\frac{dP_3(t)}{dt} = \frac{dC_3(t)}{dt} e^{-\lambda t} - C_3(t) \lambda e^{-\lambda t}. \quad (1.59)$$

Сопоставляя (1.57) и (1.59), будем иметь

$$\frac{dC_3(t)}{dt} e^{-\lambda t} - C_3(t) \lambda e^{-\lambda t} = -\lambda C_3(t) e^{-\lambda t} + \lambda \frac{(\lambda t)^2}{2} e^{-\lambda t}.$$

После приведения подобных членов и сокращения на экспоненту получим:

$$\frac{dC_3(t)}{dt} = \lambda \frac{(\lambda t)^2}{2}; \quad dC_3(t) = \lambda \frac{(\lambda t)^2}{2} dt; \quad dC_3(t) = \frac{\lambda^3 t^2}{2} dt; \quad C_3(t) = \frac{(\lambda t)^3}{2 \cdot 3} + c_3,$$

где c_3 — постоянная интегрирования, подлежащая определению из начальных условий.

Полученное решение относительно $C_3(t)$ подставим в (1.58):

$$P_3(t) = \left(\frac{(\lambda t)^3}{2 \cdot 3} + c_3 \right) e^{-\lambda t} = \frac{(\lambda t)^3}{2 \cdot 3} e^{-\lambda t} + c_3 e^{-\lambda t}.$$

Для определения постоянной интегрирования обратимся к начальным условиям (1.45) относительно $P_3(t)$:

$$t = 0; \quad P_3(0) = 0.$$

Откуда следует, что $c_3 = 0$.

Решение для $P_3(t)$ будет иметь вид

$$P_3(t) = \frac{(\lambda t)^3}{2 \cdot 3} e^{-\lambda t}$$

или, представляя знаменатель как три факториал, получим

$$P_3(t) = \frac{(\lambda t)^3}{3!} e^{-\lambda t}. \quad (1.60)$$

Таким образом, можно продолжить поиск решений и определить решение k -го дифференциального уравнения системы (1.44). Это решение по своей структуре будет аналогичным решению (1.60):

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (1.61)$$

Полученное выражение (1.61) является *распределением Пуассона*, которое определяет собой вероятность того, что на отрезке времени длиной t появится точно k требований. Для вычисления вероятности появления k требований на отрезке времени t достаточно знать интенсивность поступления требований пуассоновского потока. Распределение Пуассона является распределением вероятностей дискретной случайной величины.

1.6.2. Числовые характеристики Пуассоновского потока

Получим такие числовые характеристики Пуассоновского потока как математическое ожидание и дисперсию. Поскольку появление требований (например, k требований) — это случайное событие, то сами требования будут являться случайными дискретными величинами. Приведем определение из [11, с.38]: случайная величина, принимающая конечное или счетное число значений на числовой прямой, называется *дискретной*.

Для математического ожидания примем обозначение $M_k(t)$, где k относится к случайной величине, а t — это временной промежуток, на котором вычисляется математическое ожидание. *Математическое ожидание* — средневзвешенное значение случайной величины с весами-вероятностями [11, с.44].

Для рассматриваемого случая будем иметь

$$M_k(t) = \sum_{k=0}^{\infty} k P_k(t),$$

где $P_k(t)$ — вероятность появления k требований на отрезке времени t .

Используем распределение Пуассона (1.71):

$$\begin{aligned} M_k(t) &= \sum_{k=0}^{\infty} k P_k(t) = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} = \\ &= e^{-\lambda t} \lambda t \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!}. \end{aligned}$$

Произведем замену:

$$k - 1 = z; \quad k = 1, \Rightarrow z = 0.$$

Тогда

$$M_k(t) = \lambda t e^{-\lambda t} \sum_{z=0}^{\infty} \frac{(\lambda t)^z}{z!}. \quad (1.62)$$

Сумма, стоящая в (1.62), определяет собой экспоненту, т. е.

$$e^{\lambda t} = \sum_{z=0}^{\infty} \frac{(\lambda t)^z}{z!}. \quad (1.63)$$

Подставляя (1.63) в (1.62), получим

$$M_k(t) = \lambda t e^{-\lambda t} \sum_{z=0}^{\infty} \frac{(\lambda t)^z}{z!} = \lambda t e^{-\lambda t} e^{\lambda t} = \lambda t.$$

Таким образом, математическое ожидание — это не случайная линейная функция времени с коэффициентом пропорциональности, равным интенсивности простейшего потока. Запомним выражение для математического ожидания:

$$M_k(t) = \lambda t. \quad (1.64)$$

Можно сказать, что ожидаемое (среднее) число поступающих требований в промежутке времени $(0, t)$ равно λt .

Перейдем к вычислению дисперсии. Для дисперсии простейшего потока примем обозначение $D_k(t)$. *Дисперсия* — это математическое ожидание квадрата отклонения случайной величины от своего математического ожидания [11, с.48]. В соответствии с определением дисперсии выполним сначала следующие преобразования, в которых операция взятия математического ожидания будет обозначаться как M :

$$\begin{aligned} D_k(t) &= M[k - M(k)]^2 = M[k^2 - 2kM(k) + M^2(k)] = \\ &= M(k^2) - 2M(k)M(k) + M^2(k) = M(k^2) - M^2(k). \end{aligned} \quad (1.65)$$

Приведено известное выражение для вычисления дисперсии: дисперсия — это разность математического ожидания от квадрата случайной величины и квадрата математического ожидания случайной величины. Сначала найдем математическое ожидание от квадрата случайной величины, принимая во внимание, что квадрат случайной величины имеет то же самое распределение вероятностей:

$$\begin{aligned} M(k^2) &= \sum_{k=0}^{\infty} k^2 P_k(t) = \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} e^{-\lambda t} = e^{-\lambda t} \sum_{k=1}^{\infty} k^2 \frac{(\lambda t)^k}{k!} = e^{-\lambda t} \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{(k-1)!} = \\ &= e^{-\lambda t} \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{(k-1)!} = e^{-\lambda t} \lambda t \sum_{k=1}^{\infty} k \frac{(\lambda t)^{k-1}}{(k-1)!}. \end{aligned}$$

Сделаем замену переменной суммирования:

$$k - 1 = z; \Rightarrow k = z + 1, \quad k = 1, \quad z = 0.$$

Продолжим преобразования с новой переменной суммирования

$$M(k^2) = e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} (z+1) \frac{(\lambda t)^z}{z!}.$$

Последнее выражение представим в виде двух сумм:

$$M(k^2) = e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} (z+1) \frac{(\lambda t)^z}{z!} = e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} z \frac{(\lambda t)^z}{z!} + e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} \frac{(\lambda t)^z}{z!}.$$

Вторая сумма — это по определению экспонента с показателем λt . Тогда

$$\begin{aligned} M(k^2) &= e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} z \frac{(\lambda t)^z}{z!} + e^{-\lambda t} \lambda t \sum_{z=0}^{\infty} \frac{(\lambda t)^z}{z!} = e^{-\lambda t} \lambda t \sum_{z=1}^{\infty} z \frac{(\lambda t)^z}{z!} + e^{-\lambda t} \lambda t e^{\lambda t} = \\ &= e^{-\lambda t} \lambda t \sum_{z=1}^{\infty} \frac{(\lambda t)^z}{(z-1)!} + \lambda t = e^{-\lambda t} (\lambda t)^2 \sum_{z=1}^{\infty} \frac{(\lambda t)^{z-1}}{(z-1)!} + \lambda t. \end{aligned}$$

Произведем еще одну замену переменной суммирования:

$$z - 1 = q; \quad z = 1, \quad q = 0.$$

После подстановки новой переменной суммирования получим

$$M(k^2) = e^{-\lambda t} (\lambda t)^2 \sum_{q=0}^{\infty} \frac{(\lambda t)^q}{q!} + \lambda t = e^{-\lambda t} (\lambda t)^2 e^{\lambda t} + \lambda t = (\lambda t)^2 + \lambda t,$$

где сумма была заменена экспонентой.

В соответствии с (1.75) находим дисперсию, учитывая, что квадрат математического ожидания равен $(\lambda t)^2$:

$$D_k(t) = M(k^2) - M^2(k) = (\lambda t)^2 + \lambda t - (\lambda t)^2 = \lambda t.$$

Выпишем отдельно найденное значение дисперсии:

$$D_k(t) = \lambda t. \tag{1.66}$$

Таким образом, для пуассоновского (простейшего) потока математическое ожидание и дисперсия одинаковы и равны λt .

1.6.3. Свойство вероятностей распределения Пуассона

Проанализируем распределение Пуассона

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (1.67)$$

Напомним, что $P_k(t)$ — это вероятность появления точно k требований на отрезке времени t .

На рис. 1.1 приведено семейство распределения Пуассона при интенсивности, равной 0,5 условных единиц и для k от 0 до 6, где k — число требований.

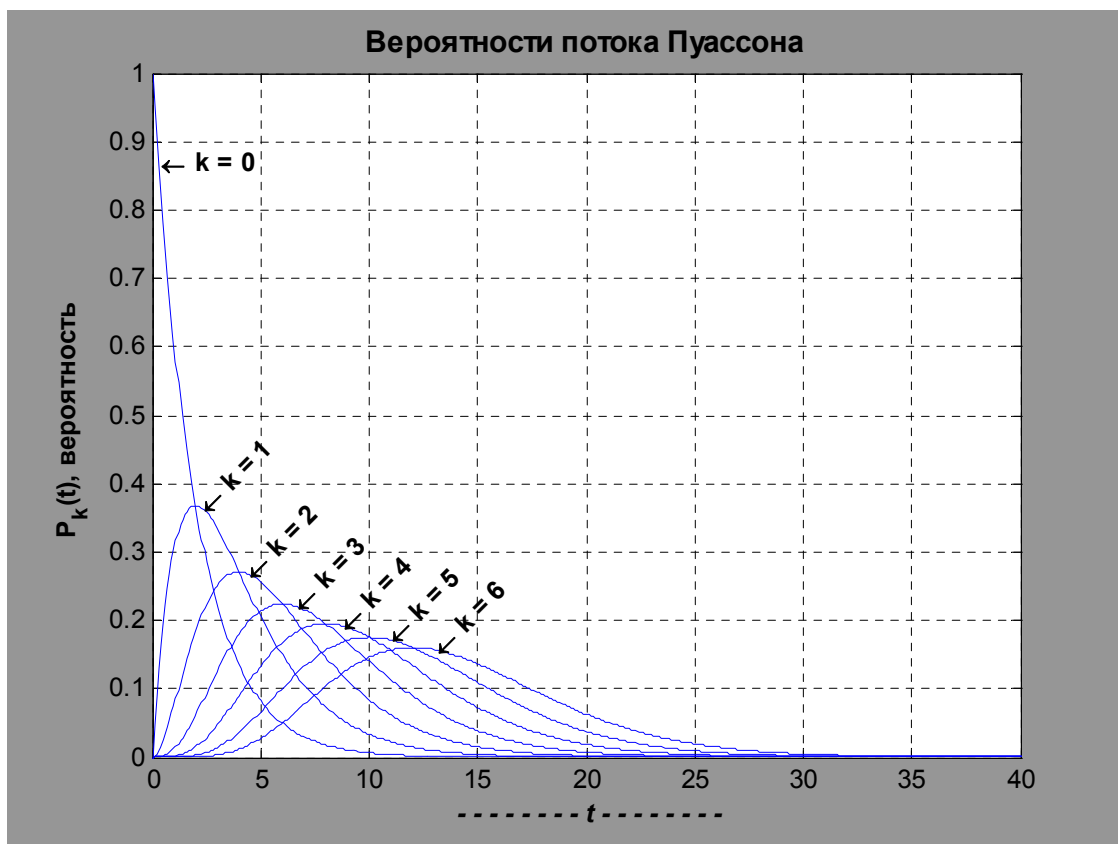


Рис. 1.1. Типовое распределение Пуассона

Исследуем распределение Пуассона на экстремум. Для этого продифференцируем (1.67) по времени t :

$$\frac{dP_k(t)}{dt} = \frac{k\lambda(\lambda t)^{k-1}}{k!} e^{-\lambda t} - \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t}. \quad (1.68)$$

Применяя необходимое условие экстремума функции, приравняем к нулю производную (1.68):

$$\begin{aligned} \frac{dP_k(t)}{dt} &= \frac{k\lambda(\lambda t)^{k-1}}{k!} e^{-\lambda t} - \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t} = 0, \\ \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} - \frac{(\lambda t)^k}{k!} e^{-\lambda t} &= 0, \\ \frac{(\lambda t)^{k-1}}{(k-1)!} - \frac{(\lambda t)^k}{k!} &= 0, \\ \frac{k(\lambda t)^{k-1} - (\lambda t)^k}{k!} &= 0, \\ k(\lambda t)^{k-1} - (\lambda t)^k &= 0, \\ k - \lambda t = 0, \Rightarrow t_3 &= \frac{k}{\lambda}. \end{aligned}$$

Точка, где функция $P_k(t)$ достигает экстремума, определяется выражением

$$t_3 = \frac{k}{\lambda}. \quad (1.69)$$

Легко проверить, что знак второй производной функции $P_k(t)$ в точке экстремума (1.69) будет отрицательным. Следовательно, функция $P_k(t)$ в точке экстремума достигает максимума. Это можно также видеть на рис. 1.1. Экстремальные точки с параметром k образуют арифметическую прогрессию:

$$\{t_{3k}\} \Rightarrow 0, \frac{1}{\lambda}, \frac{2}{\lambda}, \frac{3}{\lambda}, \dots \quad (1.70)$$

Точки ряда (1.70) равномерно расположены на числовой оси.

Вычислим значение $P_k(t)$ в точке максимума:

$$P_k(t_9) = \frac{\left(\lambda \frac{k}{\lambda}\right)^k}{k!} e^{-\lambda \frac{k}{\lambda}} = \frac{k^k}{k!} e^{-k}. \quad (1.71)$$

Максимальное значение вероятности $P_k(t)$ не зависит от интенсивности или параметра потока. Покажем, что максимумы функции $P_k(t)$ убывают с ростом k . Для этого сравним два соседних максимума функции по формуле (1.71):

$$\frac{\max P_{k+1}(t_9)}{\max P_k(t_9)} = \frac{(k+1)^{(k+1)} k! e^{-(k+1)}}{(k+1)! k^k e^{-k}} = \frac{(k+1)^k e^{-1}}{k^k} = \left(1 + \frac{1}{k}\right)^k e^{-1} < 1. \quad (1.72)$$

Соотношение (1.72) справедливо при любом конечном k , так как только при $k \rightarrow \infty$ выражение в скобках стремиться к числу e :

$$\lim_{k \rightarrow \infty} \left(1 + \frac{1}{k}\right)^k = e,$$

где $e \approx 2.71828182845905$.

Семейство кривых распределения Пуассона должно удовлетворять нормировочному условию (1.46), график которого будет располагаться параллельно горизонтальной оси на высоте, равной единице. При конечных значениях k нормировочное условие полностью не выполняется. Только при $k \rightarrow \infty$ суммарный график всего бесконечного семейства графиков будет представлять собой прямую линию, проходящую на высоте, равной единице. На рис. 1.2 приведен такой суммарный график вероятностей распределения Пуассона на отрезке времени $t \in [0, 100]$ с суммарным числом кривых, равных 20. Видно, что суммарный график располагается на высоте, равной единице только на конечном участке, а именно на отрезке времени, примерно равном 20. Далее он спадает практически до нуля, нормировочное условие не выполняется. Но на начальном отрезке времени можно считать, что нормировочное условие для поступающих из потока требований

выполняется: в каждый момент времени сумма ординат распределения Пуассона равна единице.

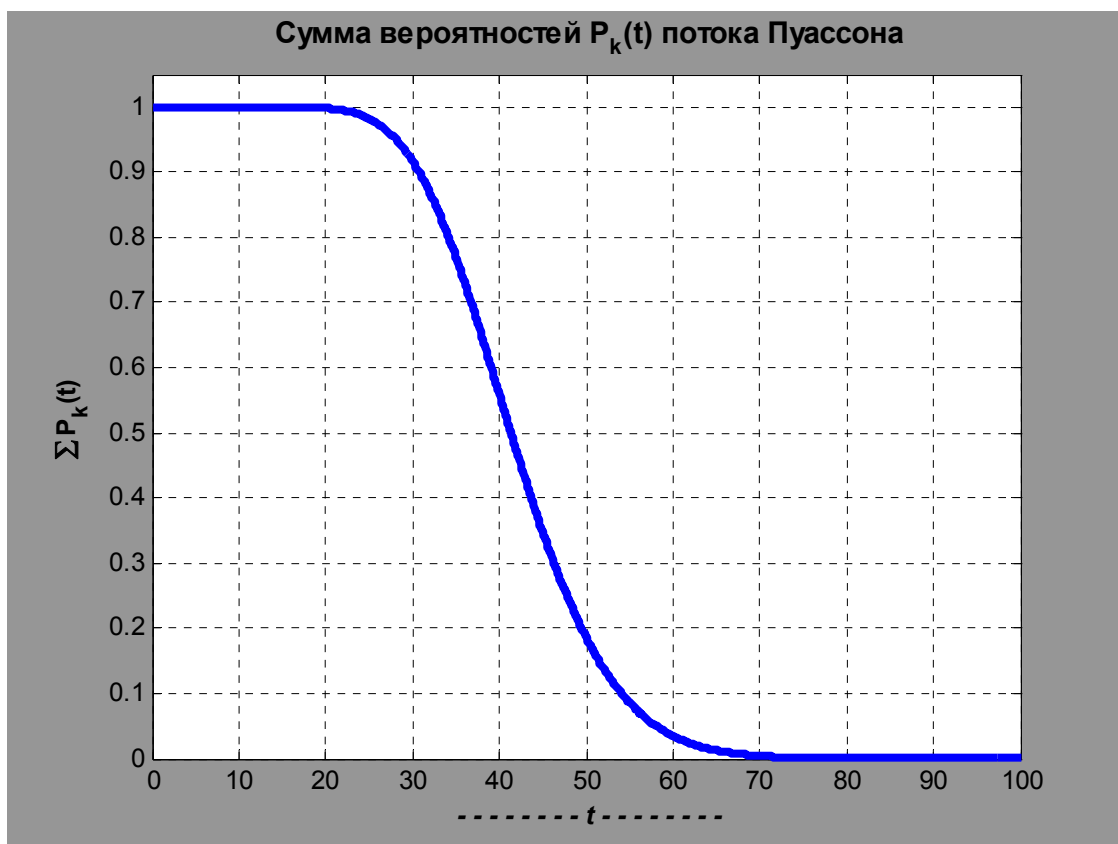


Рис. 1.2. Суммарный график вероятностей распределения Пуассона

1.6.4. Распределение длительностей интервалов между требованиями в пуассоновском потоке

Появление требований в пуассоновском потоке носит случайный характер. Очевидно, что и интервалы времени между появлениями требований будут иметь случайную длительность. Покажем, что пуассоновский поток требований характеризуется экспоненциальным (показательным) распределением промежутков времени между требованиями [4]. Время между соседними требованиями в пуассоновском потоке обозначим через τ . Величина τ является случайной

непрерывной величиной. В соответствии с определением представим функцию распределения в виде

$$F(t) = P(\tau < t), \quad (1.73)$$

где $P(\tau < t)$ — вероятность того, что случайная величина τ примет значение, меньшее t . Рассмотрим вероятность $P(\tau \geq t)$ противоположного события, т. е. вероятность того, что на участке времени длиной t , начинающемся в момент t_k , не появится ни одно из последующих требований [9]. В этом случае функция распределения (1.73) имеет вид

$$F(t) = 1 - P(\tau \geq t). \quad (1.74)$$

Величина $P(\tau \geq t)$ является вероятностью того, что в промежутке времени (t_k, t) не поступит ни одного требования. Для пуассоновского потока вероятность отсутствия требований на заданном промежутке времени определяется как $P_0(t) = e^{-\lambda t}$, где λ — интенсивность потока. При этом начало отсчета t_k можно начать с нуля, т. е. положить $t_k = 0$, поскольку пуассоновский (простейший) поток обладает свойством стационарности. Следовательно, выражение для функции распределения промежутков времени между соседними требованиями в пуассоновском потоке будет иметь вид

$$F(t) = 1 - e^{-\lambda t} \quad (t \geq 0). \quad (1.75)$$

Дифференцируя (1.75) по t , получим плотность распределения $f(t)$ промежутков времени между соседними требованиями пуассоновского потока:

$$f(t) = \lambda e^{-\lambda t} \quad (t \geq 0). \quad (1.76)$$

Полученные выражения (1.75), (1.76) описывают хорошо известное экспоненциальное (показательное) распределение [7,8,9,10,11]. Выражения (1.75), (1.76) показывают, что для пуассоновского (простейшего) потока требований промежутки времени между

соседними требованиями распределены по экспоненциальному закону с параметром λ .

Характерный вид функции плотности и функции распределения экспоненциального закона показан на рис. 1.2а.

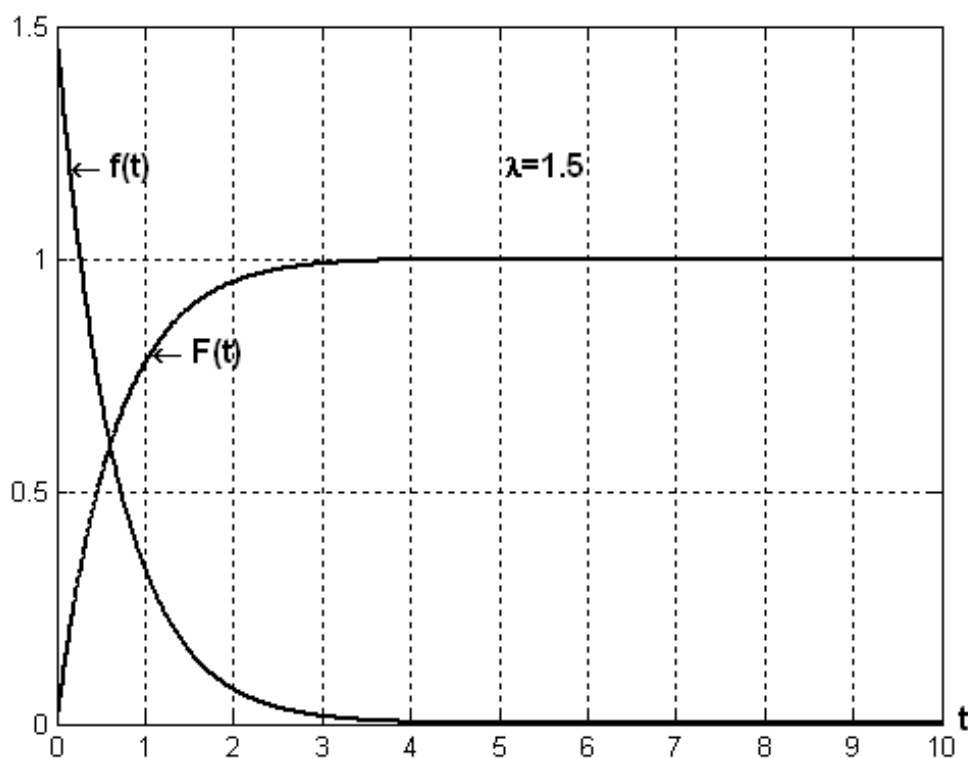


Рис. 1.2а. Функция плотности и функция распределения
экспоненциального
закона распределения длительностей интервалов времени

1.6.5. Числовые характеристики экспоненциального распределения

Найдем математическое ожидание и дисперсию длины интервала между соседними требованиями пуассоновского потока. Так как длина интервала τ между соседними требованиями является случайной

непрерывной величиной, то для вычисления математического ожидания $M(\tau)$ такой случайной величины используем следующую формулу:

$$M(\tau) = \int_0^{\infty} t \lambda e^{-\lambda t} dt, \quad (1.77)$$

где $\lambda e^{-\lambda t}$ — плотность распределения длины интервала между соседними требованиями [11, с.56].

Дисперсия $D(\tau)$ длины интервала между соседними требованиями определяется по формуле

$$D(\tau) = M(\tau^2) - [M(\tau)]^2. \quad (1.78)$$

Вычислим математическое ожидание, применяя формулу интегрирования по частям:

$$\begin{aligned} M(\tau) &= \int_0^{\infty} t \lambda e^{-\lambda t} dt = \lambda \int_0^{\infty} t e^{-\lambda t} dt = \left\{ t = u, \quad dt = du, \quad e^{-\lambda t} dt = dv, \quad v = -\frac{1}{\lambda} e^{-\lambda t} \right\} = \\ &= \lambda \left[-\frac{t e^{-\lambda t}}{\lambda} \Big|_0^{\infty} + \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda t} dt \right] = \int_0^{\infty} e^{-\lambda t} dt = -\frac{1}{\lambda} e^{-\lambda t} \Big|_0^{\infty} = \frac{1}{\lambda}. \end{aligned}$$

Таким образом, средняя длина интервала между соседними требованиями пуассоновского потока обратно пропорциональна его интенсивности. Можно также сказать, что математическое ожидание показательной случайной величины обратно пропорционально параметру λ показательного закона:

$$M(\tau) = \frac{1}{\lambda}. \quad (1.79)$$

Вычислим дисперсию. Для этого сначала в соответствии с формулой (1.78) найдем математическое ожидание от квадрата случайной величины:

$$M(\tau^2) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \lambda \int_0^{\infty} t^2 e^{-\lambda t} dt = \left\{ t^2 = u, \quad du = 2t dt, \quad e^{-\lambda t} dt = dv, \quad v = -\frac{e^{-\lambda t}}{\lambda} \right\} =$$

$$\begin{aligned}
 &= \lambda \left[-\frac{t^2 e^{-\lambda t}}{\lambda} \Big|_0^\infty + \frac{2}{\lambda} \int_0^\infty t e^{-\lambda t} dt \right] = 2 \int_0^\infty t e^{-\lambda t} dt = \\
 &= \left\{ t = u, dt = du, e^{-\lambda t} dt = dv, v = -\frac{e^{-\lambda t}}{\lambda} \right\} = 2 \left[-\frac{t e^{-\lambda t}}{\lambda} \Big|_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda t} dt \right] = \\
 &= -\frac{2e^{-\lambda t}}{\lambda^2} \Big|_0^\infty = \frac{2}{\lambda^2}. \\
 &= \lambda \left[-\frac{t^2 e^{-\lambda t}}{\lambda} \Big|_0^\infty + \frac{2}{\lambda} \int_0^\infty t e^{-\lambda t} dt \right] = 2 \int_0^\infty t e^{-\lambda t} dt = \\
 &= \left\{ t = u, dt = du, e^{-\lambda t} dt = dv, v = -\frac{e^{-\lambda t}}{\lambda} \right\} = 2 \left[-\frac{t e^{-\lambda t}}{\lambda} \Big|_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda t} dt \right] = \\
 &= -\frac{2e^{-\lambda t}}{\lambda^2} \Big|_0^\infty = \frac{2}{\lambda^2}.
 \end{aligned}$$

Учитывая выражение для математического ожидания (1.79), найдем дисперсию $D(\tau)$:

$$D(\tau) = M(\tau^2) - [M(\tau)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Таким образом, дисперсия показательной случайной величины имеет вид

$$D(\tau) = \frac{1}{\lambda^2}. \quad (1.80)$$

Определим единицу измерения параметра экспоненциального закона. Так как математическое ожидание — это средняя длина интервала времени между соседними требованиями, то тогда математическое ожидание выражается через ту или иную единицу времени. Пусть за единицу времени взята одна секунда (с). В этом случае на основании (1.79) получаем

$$\lambda = \frac{1}{M(\tau)} = \left[\frac{1}{c} \right] = [c^{-1}].$$

Единица измерения параметра экспоненциального закона равна c^{-1} .

1.7. Моделирование пуассоновского потока

При имитационном моделировании систем массового обслуживания (СМО) необходимо формировать входной поток требований. Пуассоновский поток как входной поток требований, поступающий в СМО, может быть смоделирован через интервалы времени между соседними требованиями. Это означает, что моделировать следует экспоненциальный (показательный) закон, т. е. обеспечить генерирование случайных чисел, распределенных по экспоненциальному закону. Для этого применим метод обратных функций, или метод инверсии [3].

Метод инверсии можно применять к любому вероятностному распределению при условии, что функция распределения, скажем $F(t)$, равномерно распределена в интервале $[0, 1]$. Функция распределения экспоненциального закона в соответствии с (1.75) имеет вид

$$F(t) = 1 - e^{-\lambda t} \quad (t \geq 0).$$

Пусть R — случайное число из интервала $(0; 1)$. Полагая $F(t) = R$, получим

$$R = 1 - e^{-\lambda t}, \quad e^{-\lambda t} = 1 - R, \quad -\lambda t = \ln(1 - R), \quad t = -\frac{1}{\lambda} \ln(1 - R) = -\frac{1}{\lambda} \ln R.$$

Последний переход объясняется тем, что если R — случайное число из интервала $(0; 1)$, то и $(1 - R)$ также представляет собой случайное число из того же интервала, поэтому для удобства можно заменить $(1 - R)$ на R [3, с.327]. Таким образом, для генерирования скажем N случайных чисел,

соответствующих экспоненциальному распределению, необходимо каждое случайное число t — длительность интервала между соседними требованиями — определять по формуле

$$t_i = -\frac{1}{\lambda} \ln R_i, \quad R_i \in (0,1], \quad i = \overline{1, N}. \quad (1.81)$$

Генерирование экспоненциально распределенных случайных чисел (интервалов) достаточно просто может быть реализовано практически в любой вычислительной системе, где имеется генератор случайных (псевдослучайных) чисел из интервала $(0; 1)$.

1.8. Преобразования пуассоновских потоков

Под преобразованием пуассоновских потоков будем понимать случайное разрежение и объединение.

1.8.1. Случайное разрежение пуассоновского потока

Случайным разрежением потока с вероятностью p называется операция, в результате применения которой каждое требование потока с вероятностью p остается в нем, а с вероятностью $q = 1-p$ удаляется из него [7, с.34]. Эту операцию называют также *p -преобразованием* (рис. 1.3).

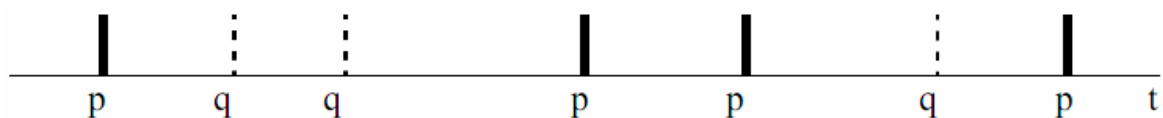


Рис. 1.3. Диаграмма разрежения потока требований

В результате p -преобразования пуассоновского потока результирующий поток будет также пуассоновским (простейшим). Для доказательства последнего утверждения следует проверить

преобразованный поток на ординарность, стационарность, отсутствие последствия. Если получающийся поток обладает этими тремя свойствами, то он будет простейшим пуассоновским потоком.

Стационарность преобразованного потока следует из стационарности исходного потока с постоянным параметром λ .

Ординарность преобразованного потока. Пусть из преобразованного потока поступает не менее 2-х требований за промежуток времени t . Тогда не менее двух требований поступают также и из исходного потока. Поскольку разреженный поток получается в результате отбрасывания какого-то количества требований из исходного потока, то вероятность $P_{\geq 2}(t)$ поступления не менее 2-х требований за время t исходного потока и вероятность $\bar{P}_{\geq 2}(t)$ преобразованного потока будут связаны следующим неравенством:

$$\bar{P}_{\geq 2}(t) \leq P_{\geq 2}(t). \quad (1.82)$$

Поделим обе части неравенства (1.82) на t и возьмем предел при $t \rightarrow 0$:

$$\lim_{t \rightarrow 0} \frac{\bar{P}_{\geq 2}(t)}{t} \leq \lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t}. \quad (1.83)$$

Для ординарного потока, каковым является пуассоновский, предел в правой части неравенства (1.83) будет равен нулю:

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} = 0. \quad (1.84)$$

Соотношение (1.84) является условием ординарности стационарного потока. Тогда из сравнения неравенства (1.83) и равенства (1.84) следует, что преобразованный поток также будет ординарным.

Отсутствие последствия разреженного потока. Для этого достаточно показать, что разреженный поток будет иметь распределение Пуассона [7]. Найдем вероятность $\bar{P}_0(t)$ отсутствия требований на отрезке времени t . В преобразованном потоке требований нет за время t в том и

только в том случае, когда их нет в исходном потоке, или в исходном потоке одно требование есть, но оно удалено с вероятностью q , или в исходном потоке есть два требования и оба удалены и т. д.:

$$\begin{aligned}\bar{P}_0(t) &= P_0(t); \\ \bar{P}_0(t) &= P_0(t) + P_1(t)q; \\ \bar{P}_0(t) &= P_0(t) + P_1(t)q + P_2(t)q^2; \\ &\dots\dots\dots,\end{aligned}$$

где $P_1(t)q$, $P_2(t)q^2, \dots$ — определены на основе испытания Бернулли, причем q, q^2, \dots — вероятности удаления требований могут быть рассчитаны по формуле Бернулли:

$$P_n(k) = C_n^k q^k p^{n-k}, \quad k = \overline{0, n},$$

где $C_n^k = \frac{n!}{k!(n-k)!}$ — биномиальные коэффициенты, n — число испытаний, k — число требуемых исходов (в рассматриваемом случае соответствует вероятности удаления требований из потока). Например, если в двух испытаниях рассчитывается вероятность успеха (т. е. вероятность удаления двух требований) в этих двух опытах, то эта вероятность будет равна

$$P_2(2) = \frac{2!}{2!(2-2)!} q^2 p^{(2-2)} = q^2.$$

Таким образом, если в пуассоновском потоке на отрезке времени t вероятность появления двух требований равна $P_2(t)$, то вероятность удаления этих двух требований равна q^2 . Так как эти события независимы между собой, то можно применить формулу произведения для двух независимых событий и рассчитать вероятность $P_2(t)q^2$, которая входит составной частью в расчет вероятности $\bar{P}_0(t)$ — вероятности отсутствия требований на отрезке времени t . После

перебора всех возможных исходов, в результате которых можно рассчитать $\bar{P}_0(t)$, применим формулу сложения вероятностей:

$$\bar{P}_0(t) = \sum_{k=0}^{\infty} P_k(t) q^k = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} q^k = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t q)^k}{k!}.$$

Заменяя сумму экспонентой с показателем $\lambda q t$, получим

$$\bar{P}_0(t) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t q)^k}{k!} = e^{-\lambda t} e^{\lambda q t} = e^{-\lambda t} e^{\lambda(1-p)t} = e^{-\lambda t} e^{\lambda t} e^{-\lambda p t} = e^{-\lambda p t}. \quad (1.85)$$

Таким образом, вероятность отсутствия требований $\bar{P}_0(t)$ в преобразованном потоке на отрезке времени t с точностью до обозначений совпадает с вероятностью отсутствия требований для исходного потока. Если вероятность отсутствия требований выражается формулой (1.85), то найдем функцию распределения $\bar{F}(t)$ интервалов времени между соседними требованиями преобразованного потока. Для этого определим функцию распределения в виде

$$\bar{F}(t) = \bar{P}(\tau < t), \quad (1.86)$$

где $\bar{P}(\tau < t)$ — вероятность того, что случайная величина τ — интервал времени между соседними требованиями — примет значение, меньшее t . Рассмотрим противоположную вероятность $\bar{P}(\tau \geq t)$, т. е. вероятность того, что на участке времени длиной t , начинающемся в момент t_k появления одного из требований потока, не появится ни одного из последующих требований [9]. Тогда функция распределения (1.73) имеет вид

$$\bar{F}(t) = 1 - \bar{P}(\tau \geq t). \quad (1.87)$$

Величина $\bar{P}(\tau \geq t)$ является вероятностью того, что в промежутке времени $(t_k; t)$ не поступит ни одного требования. Для рассматриваемого потока вероятность отсутствия требований на заданном промежутке

времени определяется как $\bar{P}_0(t) = e^{-\lambda p t}$, где (λp) — параметр потока. При этом начало отсчета t_k можно начать с нуля, т. е. положить $t_k = 0$. Следовательно, выражение для функции распределения промежутков времени между соседними требованиями в разреженном потоке будет иметь вид

$$\bar{F}(t) = 1 - e^{-\lambda p t} \quad (t \geq 0). \quad (1.88)$$

Таким образом, получили, что интервалы времени в разреженном потоке имеют экспоненциальное распределение. В свою очередь, хорошо известно, что экспоненциальное распределение обладает свойством отсутствия последействия [4]. Более того, экспоненциальное распределение промежутков времени между соседними требованиями соответствует пуассоновскому потоку, который по определению является потоком без последействия. Следовательно, преобразованный поток обладает свойством отсутствия последействия.

1.8.2. Объединение пуассоновских потоков

Считаем, что происходит объединение двух независимых пуассоновских потоков с параметрами λ_1 и λ_2 . Покажем, что если исходные потоки пуассоновские и независимы друг от друга, то их объединение также будет представлять пуассоновский поток. Для доказательства надо проверить результирующий поток на стационарность, ординарность и отсутствие последействия [7].

Стационарность результирующего потока следует из стационарности исходных потоков, поскольку у них параметры λ_1 и λ_2 постоянны.

Ординарность результирующего потока. Если предположить, что из объединенного потока за время t поступает не менее 2-х требований,

то это значит, что либо не менее 2-х поступает из одного какого-либо потока, либо из каждого поступает ровно по одному [7]. Вероятности 1-го потока будем помечать одной чертой, вероятности 2-го потока — двумя чертами, вероятности объединенного потока — без пометок. В силу независимости исходных потоков по формуле полной вероятности выразим вероятность $P_{\geq 2}(t)$ поступления не менее 2-х требований объединенного потока в виде

$$P_{\geq 2}(t) \leq \bar{P}_{\geq 2}(t) + \bar{\bar{P}}_{\geq 2}(t) + \bar{P}_1(t)\bar{\bar{P}}_1(t). \quad (1.89)$$

Последнее слагаемое в (1.89) составлено с учетом условной вероятности: вероятность появления одного требования 1-го потока возможно при наличии одного требования 2-го потока, и при этом условная вероятность равна своей безусловной вероятности, поэтому применяется формула произведения для двух независимых вероятностей. Обе части неравенства (1.89) поделим на t и перейдем к пределу при $t \rightarrow 0$:

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} \leq \lim_{t \rightarrow 0} \frac{\bar{P}_{\geq 2}(t)}{t} + \lim_{t \rightarrow 0} \frac{\bar{\bar{P}}_{\geq 2}(t)}{t} + \lim_{t \rightarrow 0} \frac{\bar{P}_1(t)\bar{\bar{P}}_1(t)}{t}. \quad (1.90)$$

Так как исходные потоки пуассоновские, обладающие свойством ординарности, то

$$\lim_{t \rightarrow 0} \frac{\bar{P}_{\geq 2}(t)}{t} = 0; \quad \lim_{t \rightarrow 0} \frac{\bar{\bar{P}}_{\geq 2}(t)}{t} = 0. \quad (1.91)$$

С учетом (1.91) перепишем (1.90):

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} \leq \lim_{t \rightarrow 0} \frac{\bar{P}_1(t)\bar{\bar{P}}_1(t)}{t}. \quad (1.92)$$

Вероятность появления одного требования для каждого из двух исходных потоков вычисляется в соответствии с распределением Пуассона:

$$\bar{P}_1(t) = \frac{\lambda_1 t}{1!} e^{-\lambda_1 t}, \quad \bar{\bar{P}}_1(t) = \frac{\lambda_2 t}{1!} e^{-\lambda_2 t}.$$

Подставляя последние выражения в (1.92), получим

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} \leq \lim_{t \rightarrow 0} \frac{\lambda_1 t e^{-\lambda_1 t} \lambda_2 t e^{-\lambda_2 t}}{t} = \lambda_1 \lambda_2 \lim_{t \rightarrow 0} t e^{-(\lambda_1 + \lambda_2)t} = 0. \quad (1.93)$$

Результат вычисления предела в (1.93) очевиден, так как под знаком предела присутствует множитель t , который по условию стремиться к нулю.

Таким образом, объединенный поток является ординарным, так как условие

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} = 0$$

является условием ординарности всякого потока.

Отсутствие последствия объединенного потока. Для этого получим формулу распределения вероятностей и сравним с распределением Пуассона. Событие, состоящее в появлении k требований объединенного потока, реализуется как пересечение события появления требований из 1-го потока и события появления требований из 2-го потока. Например, если задана вероятность появления $(k-1)$ требований 1-го потока, то для второго потока должна быть задана вероятность появления одного требования. Для определения вероятности появления k требований можно применить теорему умножения вероятностей, которая применима в случае пересечения событий. Каждое последующее событие появления k требований (например, $k-2$ требований из первого потока и 2 требования из второго потока) является несовместным с предыдущим и любым другим. Это значит, что можно применить теорему сложения вероятностей несовместных событий. Для вычисления вероятности появления k требований объединенного потока можно применить формулу полной вероятности, а именно:

$$P_k(t) = \sum_{j=0}^k \bar{P}_j(t) \bar{\bar{P}}_{k-j}(t).$$

Стоящие под знаком суммы вероятности вычисляются согласно распределению Пуассона. Следовательно,

$$\begin{aligned} P_k(t) &= \sum_{j=0}^k \bar{P}_j(t) \bar{\bar{P}}_{k-j}(t) = \sum_{j=0}^k \frac{(\lambda_1 t)^j}{j!} e^{-\lambda_1 t} \frac{(\lambda_2 t)^{k-j}}{(k-j)!} e^{-\lambda_2 t} = \\ &= e^{-(\lambda_1 + \lambda_2)t} \sum_{j=0}^k \frac{(\lambda_1 t)^j (\lambda_2 t)^{k-j}}{j! (k-j)!} = e^{-(\lambda_1 + \lambda_2)t} \sum_{j=0}^k \frac{\lambda_1^j t^j \lambda_2^k t^k \lambda_2^{-j} t^{-j}}{j! (k-j)!} = \\ &= t^k e^{-(\lambda_1 + \lambda_2)t} \sum_{j=0}^k \frac{\lambda_1^j \lambda_2^{(k-j)}}{j! (k-j)!}. \end{aligned}$$

Полученное выражение умножим и поделим на $k!$:

$$P_k(t) = t^k e^{-(\lambda_1 + \lambda_2)t} \sum_{j=0}^k \frac{\lambda_1^j \lambda_2^{(k-j)}}{j! (k-j)!} = \frac{t^k}{k!} e^{-(\lambda_1 + \lambda_2)t} \sum_{j=0}^k \frac{k! \lambda_1^j \lambda_2^{(k-j)}}{j! (k-j)!}. \quad (1.94)$$

В (1.94) последняя сумма — это бином Ньютона. Напомним его:

$$(a + b)^k = \sum_{j=0}^k \frac{k! a^j b^{(k-j)}}{j! (k-j)!}.$$

Следовательно, выражение (1.94) можно представить в виде

$$P_k(t) = \frac{[(\lambda_1 + \lambda_2)t]^k}{k!} e^{-(\lambda_1 + \lambda_2)t}. \quad (1.95)$$

Полагая $\lambda = \lambda_1 + \lambda_2$, получаем распределение Пуассона.

Таким образом, распределение вероятностей в объединенном потоке подчиняется закону Пуассона с параметром $\lambda = \lambda_1 + \lambda_2$. А поток с распределением Пуассона является потоком с отсутствием последствия. Подобный вывод можно распространить на случай объединения потоков количеством больше двух.

Сделаем еще следующее замечание. Поток требований с распределением Пуассона или простейший поток в теории массового обслуживания играет такую же роль, как нормальный закон распределения случайных величин в теории вероятностей: при сложении нескольких независимых, ординарных, случайных стационарных потоков

образуется суммарный поток, приближающийся по своим свойствам к простейшему. При этом, если система массового обслуживания обеспечивает желаемую эффективность функционирования системы при простейшем потоке требований на входе, то обслуживание системой других случайных потоков требований с одинаковой интенсивностью будет выполняться не хуже.

1.9. Потоки Эрланга

1.9.1. Определение потока Эрланга

Потоком Эрланга k -го порядка называется ординарный стационарный поток с ограниченным последствием, у которого промежутки времени T между событиями в потоке являются суммами случайных величин, распределенных по экспоненциальному закону с параметром λ [13, с.179].

Поток Эрланга k -го порядка получается из пуассоновского потока, если в последнем сохранить каждое $(k+1)$ требование при фиксированном k , а остальные отбросить [9]. Поток Эрланга получается в результате просеивания пуассоновского потока. Только в потоке Эрланга просеивание требований производится не по вероятностному закону, а строго по детерминированному, именно, когда заранее задается число отбрасываемых требований.

Операция просеивания пуассоновских потоков показана на рис. 1.4 для случая формирования потока Эрланга 2-го порядка, когда сохраняется каждое 3-требование, а два промежуточных выбрасываются.



Рис. 1.4. Формирование потока Эрланга 2-го порядка

Потоки Эрланга относятся к потокам с ограниченным последствием. Потоком с ограниченным последствием называют всякий поток, у которого промежутки времени между требованиями (событиями) в потоке образуют последовательность независимых случайных величин [13]. Стационарный ординарный поток с ограниченным последствием называется также потоком Пальма [8, 9, 13]. В потоке Пальма интервалы между требованиями (событиями) представляют собой последовательность независимых одинаково распределенных случайных величин, причем распределение может быть самым различным. В потоках Эрланга интервалы между требованиями (событиями) являются независимыми одинаково распределенными величинами, получаемые в результате суммирования интервалов времени, распределенных по экспоненциальному закону. Поток Эрланга является частным случаем потока Пальма.

1.9.2. Вывод закона распределения для интервалов времени между соседними требованиями в потоке Эрланга k -го порядка

В соответствии с определением потока Эрланга k -го порядка интервал времени T между соседними требованиями представляет собой сумму $k+1$ независимых случайных интервалов времени τ_i простейшего потока:

$$T = \sum_{i=1}^{k+1} \tau_i, \quad (1.96)$$

где τ_i — независимые случайные величины, подчиненные одному и тому же показательному закону с параметром λ .

Для показательного закона плотность вероятностей $f(t)$ определяется как

$$f(t) = \lambda e^{-\lambda t} \quad (t \geq 0). \quad (1.97)$$

Введем в рассмотрение $f_k(t)$, $F_k(t)$ — плотность распределения и функцию распределения времени T между соседними требованиями в потоке Эрланга k -го порядка. Связь между всякой плотностью распределения случайной непрерывной величины и ее функцией распределения определяется соотношением:

$$f_k(t) = \frac{dF_k(t)}{dt}. \quad (1.98)$$

Тогда $dF_k(t) = f_k(t)dt = P(t \leq T \leq t + dt)$ — это вероятность того, что T примет значение между t и $t + dt$. На рис. 1.5 показано попадание величины T в интервал $(t; t + dt)$, т. е. правая граница промежутка T окажется в пределах элементарного участка $(t; t + dt)$.

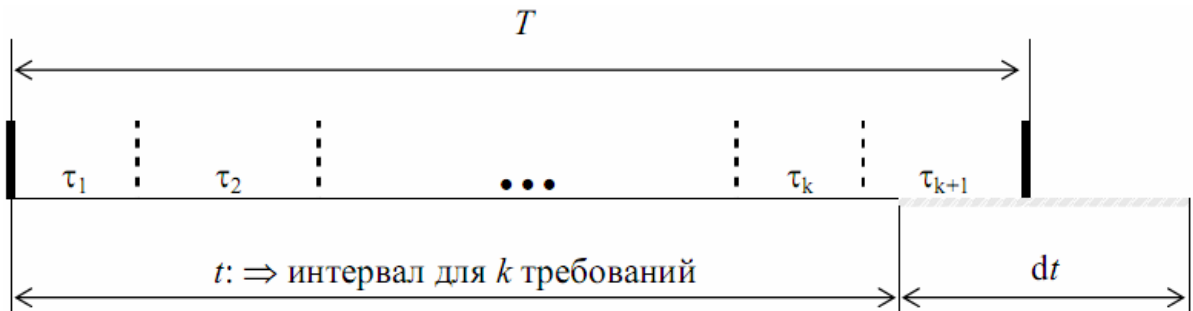


Рис. 1.5. Формирование потока Эрланга k -го порядка

Вероятность попадания k требований на отрезок времени, равный t , определяется распределением Пуассона:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (1.99)$$

Вероятность попадания одного требования на отрезок времени, равный dt , пропорциональна длине этого отрезка (см. 1.42). Поэтому можем записать

$$P_1(dt) = \lambda dt. \quad (1.100)$$

Для потока Эрланга k -го порядка попадание $k+1$ требований на отрезок

времени $[t; t + dt]$ будет определять собой случайную величину T . То, что последняя точка промежутка T попадет на отрезок $[t; t + dt]$, является вероятностной величиной, определенной выше как $P(t \leq T \leq t + dt)$. Эта вероятность определяется по формуле умножения вероятностей двух независимых событий:

$$P(t \leq T \leq t + dt) = P_k(t)P_1(dt). \quad (1.101)$$

Принимая во внимание (1.99), (1.100) и представляя вероятность $P(t \leq T \leq t + dt)$ через плотность функции распределения Эрланга k -го порядка, получим

$$P(t \leq T \leq t + dt) = f_k(t)dt = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \lambda dt,$$

откуда находим

$$f_k(t) = \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t} \quad (t \geq 0). \quad (1.102)$$

Закон распределения с плотностью (1.102) называется *законом Эрланга k -го порядка*. Формула (1.102) определяет собой плотность распределения интервалов времени t (на рис. 1.5 соответствует величине T) между соседними требованиями. Закон Эрланга может быть представлен в несколько иной форме [4,5,8,13], когда просеивание простейшего потока ведется с отсчетом требований не $k+1$, а именно k . В этом случае закон Эрланга k -го порядка будет иметь плотность распределения вероятностей вида

$$f_k(t) = \frac{\lambda(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} \quad (t \geq 0). \quad (1.103)$$

Как видно из сравнения (1.102) и (1.103) порядки потоков Эрланга отличаются на единицу. На рис. 1.6 приведены несколько реализаций функций плотности потоков Эрланга в соответствии с формулой (1.102).

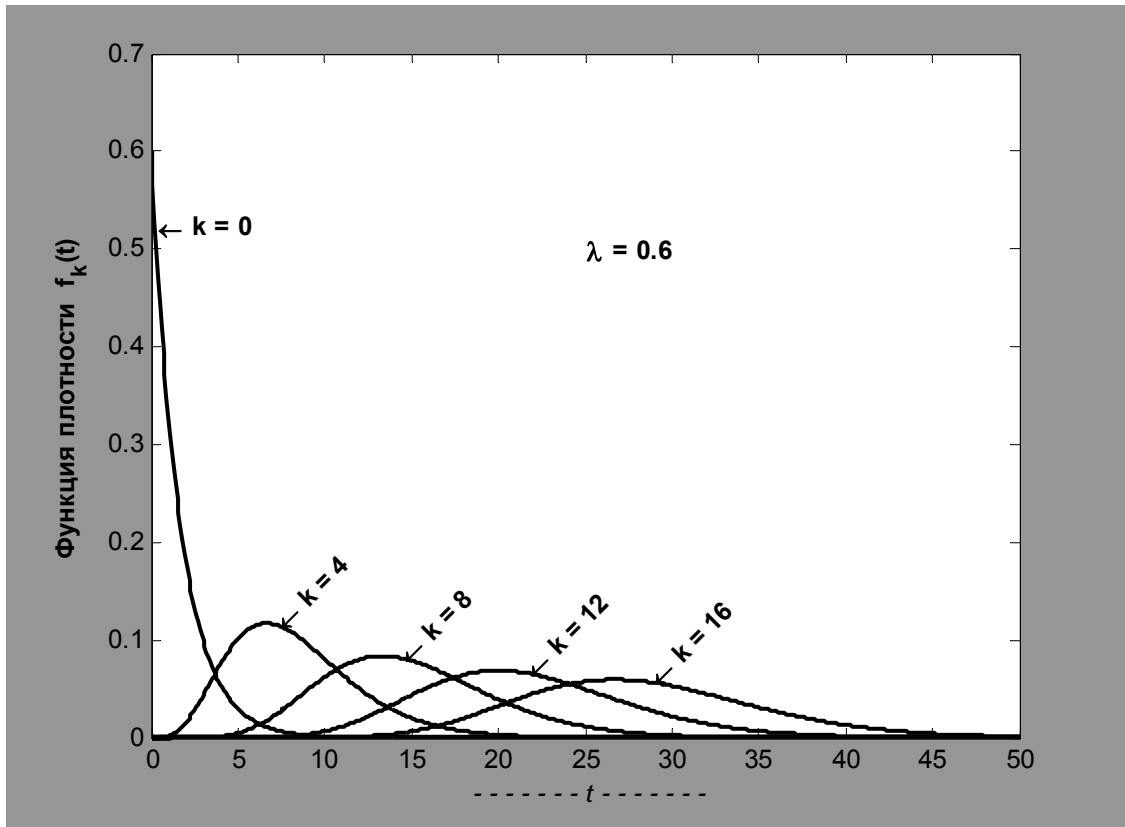


Рис. 1.6. Плотности распределения потоков Эрланга

Закон Эрланга описывает случайную непрерывную величину t — интервал времени между соседними требованиями в потоке Эрланга.

1.9.3. Функция распределения потока Эрланга

Определим функцию распределения $F_k(t)$ промежутка времени между соседними требованиями в потоке Эрланга k -го порядка. По выше приведенной формуле (1.108) определим функцию распределения $F_k(t)$ как

$$F_k(t) = \int_0^t f_k(\tau) d\tau = \int_0^t \frac{\lambda(\lambda\tau)^k}{k!} e^{-\lambda\tau} d\tau = \frac{\lambda^{(k+1)}}{k!} \int_0^t \tau^k e^{-\lambda\tau} d\tau.$$

Интеграл, стоящий в правой части равенства берется по частям, причем интегрирование по частям производится столько раз, сколько раз понадобится понизить степень первого сомножителя до нуля. Получить

общую формулу для функции распределения $F_k(t)$ можно по индукции.

Рассмотрим сначала получение $F_k(t)$ для потока Эрланга 3-го порядка:

$$\begin{aligned}
 F_3(t) &= \int_0^t f_3(\tau) d\tau = \int_0^t \frac{\lambda(\lambda\tau)^3}{3!} e^{-\lambda\tau} d\tau = \frac{\lambda^4}{3!} \int_0^t \tau^3 e^{-\lambda\tau} d\tau = \\
 &= \left\{ \tau^3 = u, \quad du = 3\tau^2 d\tau, \quad e^{-\lambda\tau} d\tau = dv, \quad v = -\frac{1}{\lambda} e^{-\lambda\tau} \right\} = \\
 &= \frac{\lambda^4}{3!} \left[-\frac{\tau^3 e^{-\lambda\tau}}{\lambda} \Big|_0^t + \frac{3}{\lambda} \int_0^t \tau^2 e^{-\lambda\tau} d\tau \right] = \frac{\lambda^4}{3!} \left[-\frac{t^3 e^{-\lambda t}}{\lambda} + \frac{3}{\lambda} \int_0^t \tau^2 e^{-\lambda\tau} d\tau \right] = \\
 &= \left\{ \tau^2 = u, \quad du = 2\tau d\tau, \quad e^{-\lambda\tau} d\tau = dv, \quad v = -\frac{1}{\lambda} e^{-\lambda\tau} \right\} = \\
 &= -\frac{(\lambda t)^3}{3!} e^{-\lambda t} + \frac{\lambda^3}{2!} \left[-\frac{\tau^2 e^{-\lambda\tau}}{\lambda} \Big|_0^t + \frac{2}{\lambda} \int_0^t \tau e^{-\lambda\tau} d\tau \right] = \\
 &= -\frac{(\lambda t)^3}{3!} e^{-\lambda t} - \frac{(\lambda t)^2}{2!} e^{-\lambda t} + \frac{\lambda^2}{1!} \int_0^t \tau e^{-\lambda\tau} d\tau = \\
 &= \left\{ \tau = u, \quad du = d\tau, \quad e^{-\lambda\tau} d\tau = dv, \quad v = -\frac{e^{-\lambda\tau}}{\lambda} \right\} = \\
 &= -\frac{(\lambda t)^3}{3!} e^{-\lambda t} - \frac{(\lambda t)^2}{2!} e^{-\lambda t} + \frac{\lambda^2}{1!} \left[-\frac{\tau e^{-\lambda\tau}}{\lambda} \Big|_0^t + \frac{1}{\lambda} \int_0^t e^{-\lambda\tau} d\tau \right] = \\
 &= -\frac{(\lambda t)^3}{3!} e^{-\lambda t} - \frac{(\lambda t)^2}{2!} e^{-\lambda t} - \frac{\lambda t}{1!} e^{-\lambda t} - \frac{\lambda}{\lambda} e^{-\lambda\tau} \Big|_0^t = \\
 &= -\frac{(\lambda t)^3}{3!} e^{-\lambda t} - \frac{(\lambda t)^2}{2!} e^{-\lambda t} - \frac{\lambda t}{1!} e^{-\lambda t} - e^{-\lambda t} + 1.
 \end{aligned}$$

Представив полученное выражение как сумму ряда, будем иметь

$$F_3(t) = 1 - \sum_{j=0}^3 \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (1.104)$$

Результат (1.104) легко можно обобщить для произвольного номера k :

$$F_k(t) = 1 - \sum_{j=0}^k \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (1.105)$$

Функция распределения (1.105) определяет собой вероятность того, что в потоке Эрланга k -го порядка интервал между соседними требованиями не превысит величины t . *Всякая функция распределения случайной величины заключена в интервале $[0; 1]$.*

1.9.4. Числовые характеристики потока Эрланга k -го порядка

Математическое ожидание M_k интервала времени между соседними требованиями в потоке Эрланга k -го порядка:

$$M_k = \sum_{i=1}^{k+1} M_i = \sum_{i=1}^{k+1} \frac{1}{\lambda} = \frac{k+1}{\lambda}, \quad (1.106)$$

где $M_i = \frac{1}{\lambda}$ — математическое ожидание показательного распределенной величины промежутка времени между соседними требованиями в исходном пуассоновском потоке. Формулу (1.106) получаем на основе свойства математического ожидания: математическое ожидание суммы независимых случайных величин равно сумме математических ожиданий этих величин. А как известно — в соответствии с определением потока Эрланга — интервал времени между соседними требованиями в потоке Эрланга равняется сумме промежутков времени между соседними требованиями в исходном пуассоновском потоке. Математическое ожидание (1.116) определяет собой среднюю величину интервала времени между соседними требованиями в потоке Эрланга k -го порядка. Математическое ожидание может быть вычислено также по общей формуле математического ожидания случайных непрерывных величин:

$$M_k = \int_0^{\infty} t f_k(t) dt.$$

Подставляя значение $f_k(t)$ в соответствии с (1.112), математическое ожидание будет определяться по выражению

$$M_k = \int_0^{\infty} t f_k(t) dt = \int_0^{\infty} t \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t} dt,$$

из которого после многократного интегрирования по частям будет получен результат (1.116).

По известному математическому ожиданию может быть вычислена интенсивность Λ_k потока Эрланга k -го порядка:

$$\Lambda_k = \frac{1}{M_k} = \frac{\lambda}{k+1}, \quad (1.107)$$

где λ — интенсивность исходного пуассоновского потока.

Дисперсия интервала времени между соседними требованиями в потоке Эрланга может быть определена через сумму дисперсий промежутков времени между соседними требованиями в исходном пуассоновском потоке. Промежутки времени в исходном пуассоновском потоке — это независимые случайные величины с показательным законом распределения. Поэтому справедливо свойство дисперсии суммы независимых случайных величин [11]: дисперсия суммы независимых случайных величин равна сумме дисперсий. Для показательно распределенной случайной величины ее дисперсия определяется формулой (1.80). Тогда искомая дисперсия D_k равна:

$$D_k = \sum_{i=1}^{k+1} D_i = \sum_{i=1}^{k+1} \frac{1}{\lambda^2} = \frac{k+1}{\lambda^2}, \quad (1.108)$$

где $D_i = \frac{1}{\lambda^2}$ — дисперсия промежутка времени между соседними

требованиями в исходном пуассоновском потоке.

В случае, когда закон Эрланга определен формулой (1.103), то числовые характеристики потока Эрланга примут следующий вид:

$$M_k = \sum_{i=1}^k M_i = \sum_{i=1}^k \frac{1}{\lambda} = \frac{k}{\lambda}, \quad (1.109)$$

$$\Lambda_k = \frac{1}{M_k} = \frac{\lambda}{k}, \quad (1.110)$$

$$D_k = \sum_{i=1}^k D_i = \sum_{i=1}^k \frac{1}{\lambda^2} = \frac{k}{\lambda^2}. \quad (1.111)$$

1.9.5. Нормированный поток Эрланга

Интенсивность Λ_k потока Эрланга может быть использована для вычисления числовых характеристик M_k , D_k :

$$M_k = \frac{k+1}{\lambda} = \frac{1}{\Lambda_k}, \quad (1.112)$$

$$D_k = \frac{k+1}{\lambda^2} = \frac{(k+1)^2}{\lambda^2(k+1)} = \frac{1}{(k+1)\Lambda_k^2}, \quad (1.113)$$

где Λ_k вычисляется по формуле (1.107).

Интенсивность Λ_k может быть введена в закон Эрланга вида (1.102):

$$f_k(t) = \frac{(k+1)\Lambda_k[(k+1)\Lambda_k t]^k}{k!} e^{-(k+1)\Lambda_k t} = \frac{[(k+1)\Lambda_k]^{(k+1)} t^k}{k!} e^{-(k+1)\Lambda_k t}. \quad (1.114)$$

Если принять, что интенсивность потока постоянна, т. е. $\Lambda_k = \text{const}$, то математическое ожидание не зависит от порядка потока k , а дисперсия с возрастанием k неограниченно убывает [13, с.181]. Это означает, что

поток Эрланга с увеличением k и постоянной интенсивности приближается к регулярному потоку с постоянными интервалами времени между требованиями. Анализ функции вида (1.114), позволяет сделать вывод, что при $k \rightarrow \infty$ функция плотности $f_k(t)$ при постоянной интенсивности обращается в δ -функцию Дирака в точке $t = 1/\Lambda_k$. Процесс обращения $f_k(t)$ в δ -функцию показан на рис. 1.7.

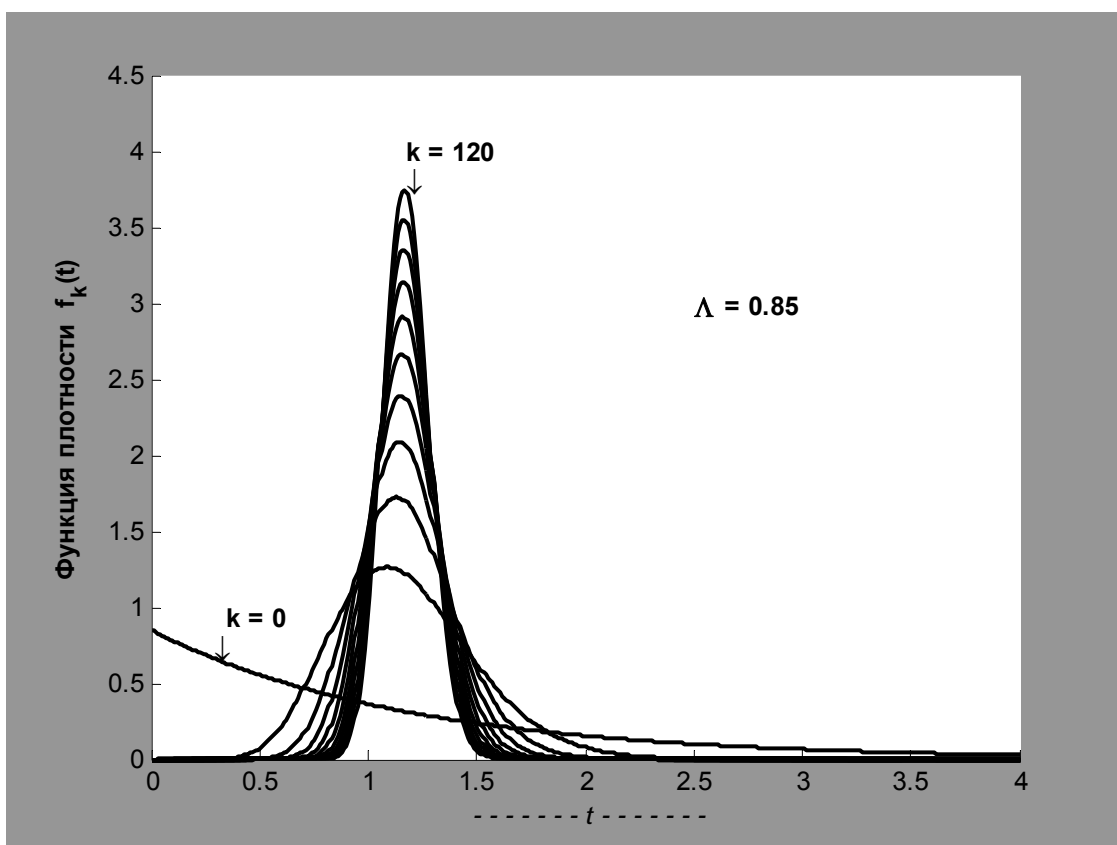


Рис. 1.7. Процесс изменения функции плотности потока Эрланга

Указанное свойство потоков Эрланга позволяет при различных k получать практически любую степень последствия потока — от полного отсутствия последствия (при $k = 0$) до жесткой функциональной связи между моментами появления требований (регулярный поток) при $k \rightarrow \infty$ [13].

Поток Эрланга с постоянной интенсивностью Λ_k можно назвать нормированным потоком.

1.9.6. Замена реальных потоков требований потоками Эрланга

Произвольные потоки с ограниченным последствием можно заменить потоками Эрланга с теми же числовыми характеристиками для промежутка времени между требованиями в потоке [13]. Процедуру замены произвольного потока потоком Эрланга рассмотрим на конкретном примере. Пусть для исследуемого потока получены математическое ожидание $m=2$ и дисперсия $d=0.8$ величины T — промежутка времени между соседними требованиями. В соответствии с формулами (1.106), (1.108) для потока Эрланга любого порядка имеем следующую связь между математическим ожиданием M_k и дисперсией D_k :

$$M_k = \frac{k+1}{\lambda}, \quad D_k = \frac{k+1}{\lambda^2}, \quad M_k = \lambda D_k, \quad \lambda = \frac{M_k}{D_k}.$$

Последнее выражение используем для определения интенсивности потока Эрланга по числовым характеристикам заменяемого потока:

$$\lambda = \frac{m}{d} = \frac{2}{0.8} = 2,5.$$

Из равенств $m = \frac{k+1}{\lambda}$ и $d = \frac{k+1}{\lambda^2}$, получаем

$$k+1 = m\lambda = 2 \cdot 2,5 = 5;$$

$$k+1 = d\lambda^2 = 0,8 \cdot 2,5^2 = 0,8 \cdot 6,25 = 5.$$

По любому из приведенных равенств получаем порядок k потока Эрланга:

$$k+1 = 5; \quad k = 5 - 1 = 4.$$

Можно также определить порядок потока Эрланга через числовые характеристики:

$$k+1 = m\lambda = \frac{m^2}{d}, \quad k = \frac{m^2}{d} - 1 = \frac{2^2}{0.8} - 1 = 5 - 1 = 4.$$

Таким образом, рассматриваемый поток с ограниченным последствием с известными числовыми характеристиками промежутка времени между соседними требованиями можно заменить потоком Эрланга 4-го порядка с интенсивностью $\lambda = 2,5$.

Рассмотренная достаточно простая процедура замены имеет некоторые ограничения [13]: во-первых, замена возможна при соблюдении условия

$$d \leq m^2,$$

где d , m — дисперсия и математическое ожидание промежутка времени между соседними требованиями в заменяемом потоке. Кроме того, функциональная зависимость математического ожидания и дисперсии промежутка времени между соседними требованиями в потоке Эрланга предполагает, что получаемый порядок потока Эрланга должен выражаться только натуральным числом k . В случае дробного числа k замена возможна так называемым обобщенным потоком Эрланга, который получается из пуассоновского с различными параметрами. В этом случае величина T — интервал времени между соседними требованиями будет представлять сумму случайных величин, которые равны промежуткам времени между соседними требованиями пуассоновских потоков с различными параметрами. Обобщенный закон Эрланга можно будет получить в результате композиции двух или нескольких пуассоновских потоков с различными параметрами [8,9,10,11].

1.9.7. Моделирование потока Эрланга

В задачу моделирования потока Эрланга входит формирование интервалов времени T_i между соседними требованиями. А это означает, что необходимо получение последовательности чисел, определяемых эрланговским распределением порядка k [14]. Интервалы времени T_i есть случайные числа. Чтобы получить случайное число, распределенное по закону Эрланга порядка k , его нужно вычислить как сумму k экспоненциально распределенных случайных величин [14, с.242]. Формирование экспоненциально распределенных чисел осуществляется по формуле (1.91):

$$t_i = -\frac{1}{\lambda} \ln R_i, \quad R_i \in (0; 1], \quad i = \overline{1, N},$$

где N — желаемое число генерируемых чисел.

Для формирования одного числа T с эрланговским распределением k -го порядка необходимо просуммировать полученные числа с экспоненциальным распределением, т. е.

$$T = -\frac{1}{\lambda} \sum_{i=1}^k \ln(R_i), \quad R_i \in (0; 1], \quad (1.115)$$

где k — порядок распределения Эрланга. В последнем выражении логарифм вычисляется k раз. Используя свойства логарифмов, сумму логарифмов заменим на логарифм произведения. Тогда вычисление одного числа с эрланговским распределением можно производить по формуле

$$T = -\frac{1}{\lambda} \ln \left(\prod_{i=1}^k R_i \right), \quad R_i \in (0; 1]. \quad (1.116)$$

Формулы (1.115), (1.116) математически эквивалентны. Только при машинном вычислении формула (1.116) предпочтительнее, так как в ней

логарифм вычисляется один раз, поэтому вычисление по формуле (1.116) в ЭВМ будет происходить быстрее, чем вычисление по формуле (1.115), принципиально, во всяком случае [14].

Для получения N чисел с эрланговским распределением необходимо производить циклическое вычисление величины T по одной из формул (1.115), (1.116). Таким образом, для формирования чисел с эрланговским распределением необходимо иметь в распоряжении датчик случайных чисел, равномерно распределенных в интервале $(0; 1]$. Современные машинные вычислительные системы снабжены датчиками равномерно распределенных случайных (псевдослучайных) чисел. Поэтому создать программу для вычисления случайных чисел с эрланговским распределением по одной из формул (1.115), (1.116) не должно вызывать затруднений.

Кроме того, бывает необходимо рассчитать вероятность того или иного количества требований в потоке Эрланга. Для этого можно воспользоваться потоком Пуассона, в котором интенсивность потока будет уменьшена в $(k+1)$ раз, т. е. распределение вероятностей в потоке Эрланга r -го порядка будет определяться по следующей формуле:

$$P_{kr}(t) = \frac{((\lambda/(r+1))t)^k}{k!} e^{-(\lambda/(r+1))t},$$

где $P_{kr}(t)$ — вероятность появления точно k требований на отрезке времени длиной t в потоке Эрланга r -го порядка.

1.10. Гиперэкспоненциальное распределение

В случае, когда m параллельно работающих приборов обслуживания составляют один обобщенный прибор обслуживания, то часто обслуживается только одно требование, которое с какой-либо

вероятностью выбирается на обслуживание. Если обслуживание происходит по экспоненциальному закону, то при вероятностном выборе экспоненциально распределенной обслуживающей ветви плотность распределения вероятностей времени обслуживания будет подчиняться гиперэкспоненциальному распределению [14, с. 243]. Плотность $r(t)$ гиперэкспоненциального распределения задается соотношением

$$r(t) = \sum_{i=1}^n c_i \lambda_i e^{-\lambda_i t}, \quad t \geq 0, \quad (1.117)$$

где c_i — вероятности выбора экспоненциально распределенной ветви с интенсивностью λ_i , $\sum_{i=1}^n c_i = 1$. Гиперэкспоненциальное распределение обозначают H_R [4], где R — число экспоненциально распределенных ветвей обслуживания. В частности, H_3 означает, что гиперэкспоненциальное распределение осуществляется при 3-х ветвях. Систему с гиперэкспоненциальным обслуживанием называют системой с параллельными этапами обслуживания. На рис. 1.8 показана схема системы обслуживания с параллельно включенными ветвями (приборами, каналами), на которые с вероятностью c_i отправляется на обслуживание требование, поступившее в систему обслуживания.

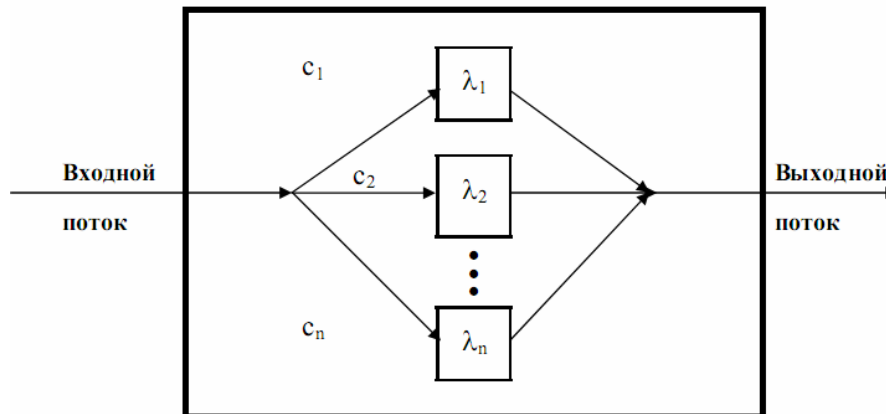


Рис. 1.8. Система с параллельными этапами обслуживания

Пример графика функции плотности гиперэкспоненциального распределения показан на рис. 1.9, где также приведены графики экспоненциального распределения.

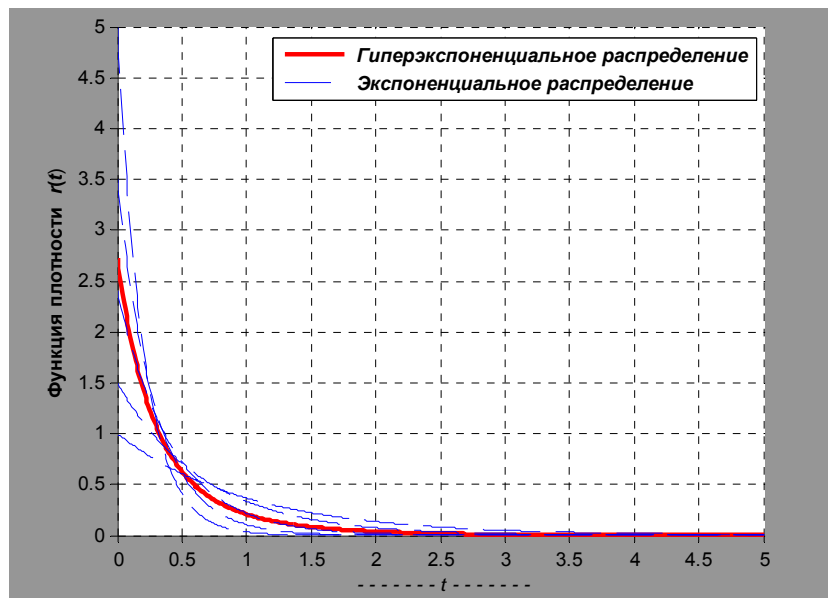


Рис. 1.9. График функции плотности гиперэкспоненциального распределения

Построение на рис. 1.9 выполнено при следующих значениях вероятностей выбора c_i и интенсивностей экспоненциального распределения:

$$c_1 = 0,3; \lambda_1 = 2,4; c_2 = 0,15; \lambda_2 = 3,5; c_3 = 0,25; \lambda_3 = 1,5; \\ c_4 = 0,2; \lambda_4 = 5; c_5 = 0,1; \lambda_5 = 1.$$

1.11. Гиперэрланговское распределение

В соответствии с [5,14] приведем следующее определение: гиперэрланговским распределением называется распределение с плотностью

$$p(t) = \sum_{r=0}^{\infty} c_r \frac{\lambda(\lambda t)^r}{r!} e^{-\lambda t}, \quad t \geq 0, \quad (1.118)$$

где $\lambda > 0$ — интенсивность потока, c_r — вероятности выбора одного из потоков Эрланга порядка r , причем $\sum_{r=0}^{\infty} c_r = 1$. Реально предел

суммирования конечен. Гиперэрланговское распределение — это смесь эрланговских распределений с весовыми коэффициентами, равными вероятностям выбора одного из эрланговских распределений. Например, если предел суммирования в (1.118) ограничивается числом 4, то это значит, что будет производиться суммирование потоков Эрланга нулевого, первого, второго, третьего, четвертого порядков с соответствующими весовыми коэффициентами c_1, c_2, c_3, c_4 таких,

что $\sum_{r=1}^4 c_r = 1$. На рис. 1.10 показан график функции плотности гиперэрланговского распределения и исходные функции плотности

обычного распределения Эрланга. Параметры распределения:

$$c_1 = 0,3; \quad c_2 = 0,15; \quad c_3 = 0,25; \quad c_4 = 0,2; \quad c_5 = 0,1; \quad \lambda = 4.$$

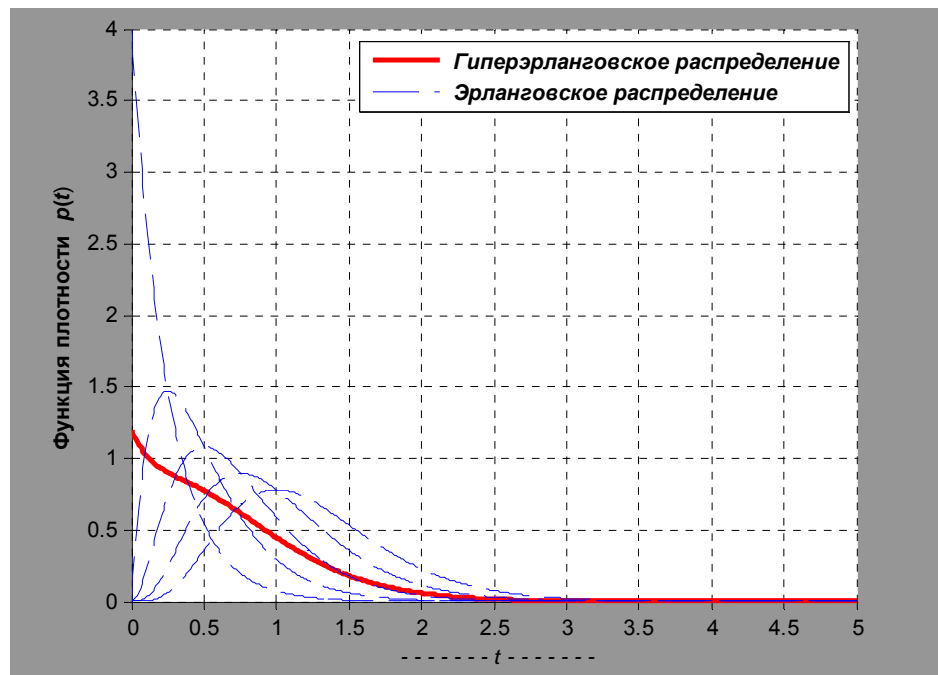


Рис. 1.10. График функции плотности гиперэрланговского распределения

Для гиперэрланговского распределения и соответственно потока можно дать следующую интерпретацию [5]. Пусть некоторая операция (например, обслуживание требования) имеет случайную длительность ξ с плотностью (1.118). Тогда можно считать, что в начале операции реализуется случайное испытание, результатом которого может быть любой исход $r = 0, 1, 2, \dots$ с вероятностью c_r . Если исход испытания — данное число r , то при этом условии ξ имеет распределение Эрланга порядка r с параметром λ .

ГЛАВА 2. ПРОЦЕССЫ РАЗМНОЖЕНИЯ И ГИБЕЛИ

2.1. Определение и постулаты процесса размножения и гибели

Термин "процесс размножения и гибели" (процесс рождения и гибели) пришел из биологии и связан с рассмотрением численности биологических объемов популяций. Говорят, что процесс находится в состоянии S_k , если объем популяции равен k ; переход из состояния S_k в состояние S_{k+1} соответствует рождению, а переход из S_k в S_{k-1} — гибели.

Процесс перехода из одного состояния в другое можно представить в виде цепи (рис. 2.1), состоящей из дискретных состояний S_i , каждое из которых связано прямой и обратной связью с двумя соседними S_{i-1} , S_{i+1} , а каждое из двух крайних связано прямой и обратной связью только с одним соседним.

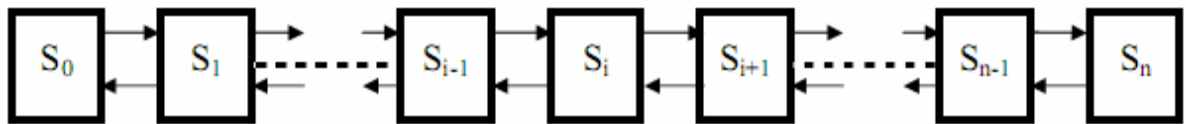


Рис. 2.1. Схема системы с дискретными состояниями

Схема, представленная на рис. 2.1, называется *схемой размножения и гибели*, а сам процесс — *процессом размножения и гибели*.

Модель системы, представленная на рис. 2.1, в виде прямоугольников (кружочков, овалов), в которые вписаны состояния системы и которые соединены между собой стрелками (дугами), называется *размеченным графом состояний*. Для процесса размножения и гибели верхняя стрелка (дуга), идущая слева направо, соответствует указанию на рождение, нижняя, идущая справа налево, указывает на гибель. Если при этом над стрелками (дугами) указано число, то оно означает интенсивность рождения или гибели. Над стрелками могут располагаться и буквенные

обозначения интенсивностей рождения и гибели. В общем случае соединения между состояниями могут быть сколь угодно различными, а не только строго из одного состояния в другое соседнее. Обозначения или числа над стрелками в общем случае означают интенсивность перехода из одного состояния в другое состояние.

Приведем еще одно определение, следующее в основном [13, с.82]: *процессом размножения и гибели* называется случайный процесс без последствия, множество состояний которого $(s_0, s_1, \dots, s_i, \dots, s_n)$ ставится в однозначное соответствие с рядом целых неотрицательных чисел $\{0, 1, 2, \dots, n\}$, и для любого состояния этого процесса (за исключением граничных состояний s_0, s_n) соседними состояниями могут быть только те, которые отличаются от индекса рассматриваемого состояния на величину ± 1 .

Процесс размножения и гибели может в некоторых случаях иметь не конечное число состояний: $s_0, s_1, \dots, s_i, \dots, s_n$, а бесконечное (счетное): $s_0, s_1, \dots, s_i, \dots$

Часто состояния процесса размножения и гибели обозначают через букву E , например, k -е состояние в новых обозначениях это E_k . Поэтому в дальнейшем будем использовать обозначения состояний как через букву S , так и через букву E .

Прежде чем перейти к постулатам процесса размножения и гибели введем следующие обозначения: λ_k — интенсивность размножения из состояния E_k или скорость, с которой происходит размножение в популяции объема k ; μ_k — интенсивность гибели из состояния E_k или скорость, с которой происходит гибель в популяции объема k . Введенные интенсивности размножения и гибели не зависят от времени, а зависят только от состояния E_k [4].

Постулаты процесса размножения и гибели [15]:

1. Изменения системы осуществляются только путем переходов из состояний в ближайшие к ним соседние состояния (из E_k в E_{k+1} или в E_{k-1} при $k \geq 1$, а из E_0 — только в E_1).
2. Если в момент времени t система находится в состоянии E_k , то вероятность того, что между t и $t + \Delta t$ произойдет переход E_k в E_{k+1} , равна $\lambda_k \Delta t + o(\Delta t)$, а вероятность перехода E_k в E_{k-1} (если $k \geq 1$) равна $\mu_k \Delta t + o(\Delta t)$. Вероятность более чем одного изменения на протяжении $(t; t + \Delta t)$ есть $o(\Delta t)$ — малая величина более высокого порядка, чем Δt .

Первый постулат отражен в определении процесса размножения и гибели, а также представлен графически на рис. 2.1. Второй постулат означает то, что процесс размножения и гибели является ординарным процессом: в системе изменение состояния на одну единицу происходит за счет рождения или за счет гибели. Во втором постулате математически отражено ординарность потока (потока рождения и потока гибели). Для процесса размножения и гибели кратные рождения, кратные гибели и одновременные рождения и гибели в течение малого промежутка времени Δt запрещены в том смысле, что вероятность таких кратных событий имеет порядок $o(\Delta t)$ [4].

Размеченный граф состояний или диаграмма интенсивностей переходов для процесса размножения и гибели с бесконечным числом состояний показан на Рис. 2.2.

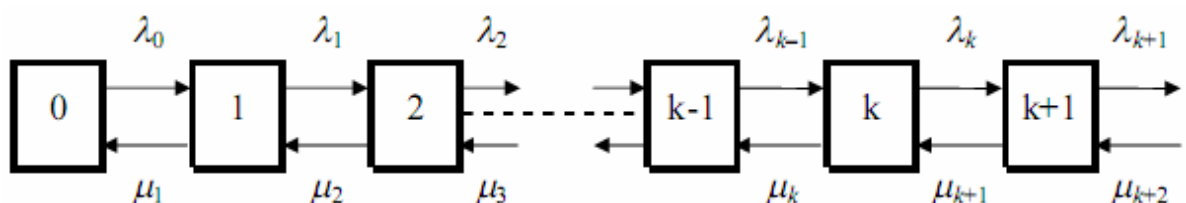


Рис. 2.2. Диаграмма интенсивностей переходов

для процесса размножения и гибели

На рис. 2.2 каждое состояние соответствует ряду натуральных чисел $0, 1, 2, \dots, k-1, k, k+1, \dots$. Из крайнего левого или нулевого состояния возможен только один переход — соответственно одно рождение — в состояние с цифрой 1 с интенсивностью λ_0 . Из любого другого ($k \geq 1$) возможно как рождение, так и гибель. Например, из k -го состояния возможно одно рождение с интенсивностью λ_k и, соответственно, переход в $k+1$ состояние, а возможен переход в $k-1$ состояние, что соответствует гибели с интенсивностью μ_k .

2.2. Дифференциальные уравнения процесса размножения и гибели

Получим дифференциальные уравнения для вероятностей $P_k(t)$ того, что система будет находиться в состоянии E_k . Для этого предварительно рассмотрим условия перехода системы в состояние E_k .

2.2.1. Условия вывода дифференциальных уравнений процесса размножения и гибели

Для вывода дифференциальных уравнений необходимо сначала вычислить вероятность того, что система будет находиться в состоянии E_k в момент времени $t + \Delta t$, т. е. $P_k(t + \Delta t)$ [4].

На момент времени $t + \Delta t$ возможны следующие исходы:

1. В момент времени t система находится в состоянии E_k и за время Δt не происходит никаких изменений. Это условие отразим в виде:

$$\begin{array}{c} \Delta t \\ E_k \longrightarrow E_k. \end{array}$$

2. В момент времени t система находится в состоянии E_{k-1} и за время Δt происходит переход в E_k . Это условие отразим в виде:

$$\begin{array}{c} \Delta t \\ E_{k-1} \longrightarrow E_k. \end{array}$$

3. В момент времени t система находится в состоянии E_{k+1} и за время Δt происходит переход в E_k . Это условие отразим в виде:

$$\begin{array}{c} \Delta t \\ E_{k+1} \longrightarrow E_k. \end{array}$$

4. Между временем t и $t + \Delta t$, т. е. за малый период Δt не может происходить более одного перехода. Это означает, что переходы в состояние E_k из состояний, не являющихся соседними, не возможны. В соответствии с постулатами размножения и гибели вероятность более одного перехода равна $o(\Delta t)$ и из рассмотрения исключается.

Если использовать биологическую терминологию, то состояние E_k соответствует объему популяции, равной k , для состояния E_{k-1} объем популяции равен $k - 1$, для состояния E_{k+1} объем популяции равен $k + 1$. При этом переход из состояния E_{k+1} в состояние E_k соответствует гибели одного члена популяции, переход из E_{k-1} в E_k соответствует рождению одного члена популяции, а в случае отсутствия изменений, когда нет ни рождения, ни гибели, то это формально соответствует переходу из E_k в E_k .

2.2.2. Определение вероятностей переходов системы в состояние E_k

Отмеченные выше возможные переходы системы в вероятностном отношении являются событиями несовместными — попарно несовместными в силу принятых постулатов. Как известно [8], несколько событий называются несовместными, если появление любого из них исключает появление каждого из остальных. Вероятности перехода системы из одного состояния в другое происходит за время Δt . Будем обозначать эти вероятности двумя нижними индексами: первый индекс указывает на исходное состояние, второй индекс — на конечное состояние. В частности, для рассматриваемой системы имеют место следующие переходные вероятности:

- вероятность рождения: $P_{k-1,k}(\Delta t) = \lambda_{k-1}\Delta t + o(\Delta t)$;
 - вероятность гибели: $P_{k+1,k}(\Delta t) = \mu_{k+1}\Delta t + o(\Delta t)$;
 - вероятность отсутствия, как рождения, так и гибели можно определить как дополнение от суммы вероятностей гибели и рождения в состоянии E_k : $P_{k,k}(\Delta t) = 1 - \lambda_k\Delta t - \mu_k\Delta t + o(\Delta t)$;
- где $\lambda_k\Delta t + o(\Delta t)$ — вероятность рождения в системе, находящейся в состоянии E_k , $\mu_k\Delta t + o(\Delta t)$ — вероятность гибели в системе, находящейся в состоянии E_k .

На момент времени t обозначим вероятности событий в виде:

$P_k(t)$ — вероятность того, что в системе на момент времени t имеется точно k членов популяции (в дальнейшем k требований),

$P_{k-1}(t)$ — вероятность того, что в системе на момент времени t имеется точно $k-1$ членов популяции,

$P_{k+1}(t)$ — вероятность того, что в системе на момент времени t имеется точно $k+1$ членов популяции.

На момент времени $t + \Delta t$ вероятность нахождения в системе k членов популяции обозначаем как $P_k(t + \Delta t)$. Расчет $P_k(t + \Delta t)$ в каждом из возможных исходов системы производится на основе теоремы о произведении вероятностей независимых событий: вероятность произведения двух независимых событий равно произведению вероятностей этих событий. В теоретико-множественном отношении теорема умножения вероятностей соответствует вероятности совместного появления независимых событий и рассматривается как пересечение этих событий. Для двух независимых событий A и B : справедливо соотношение $P(AB) = P(A) \cdot P(B)$, что можно изобразить графически (рис. 2.3) как пересечение двух множеств A и B .

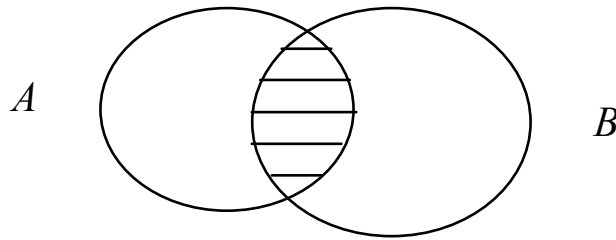
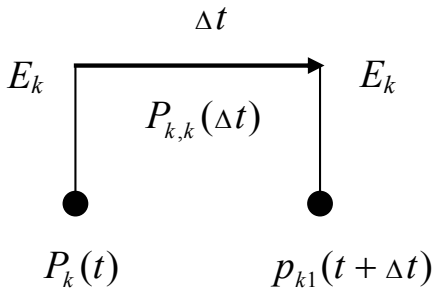


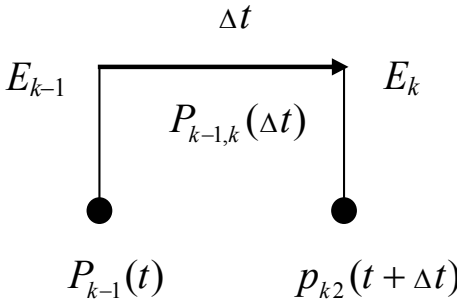
Рис. 2.3. Определение пересечения двух множеств

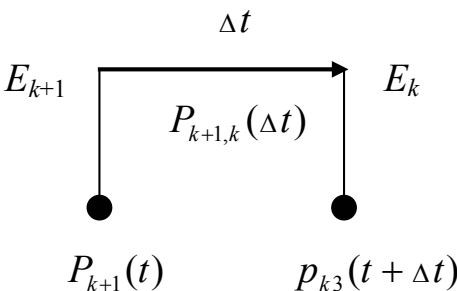
Запишем через строчную букву вероятность $p_{ki}(t + \Delta t)$, $i = \overline{1, 3}$ для каждого из переходов:

- при переходе из E_k в E_k : $p_{k1}(t + \Delta t) = P_k(t) \cdot P_{k,k}(\Delta t)$,
- при переходе из E_{k-1} в E_k : $p_{k2}(t + \Delta t) = P_{k-1}(t) \cdot P_{k-1,k}(\Delta t)$,
- при переходе из E_{k+1} в E_k : $p_{k3}(t + \Delta t) = P_{k+1}(t) \cdot P_{k+1,k}(\Delta t)$,

Дадим следующую интерпретацию для вероятностей $p_{ki}(t + \Delta t)$, $i = \overline{1, 3}$:

1. 

Переход из E_k в E_k
2. 

Переход из E_{k-1} в E_k
3. 

Переход из E_{k+1} в E_k

2.2.3. Определение вероятности состояния E_k на момент времени $t + \Delta t$

Для каждого из трех событий вероятность $p_{ki}(t + \Delta t)$, $i = 1, 2, 3$ определяется через произведения вероятностей (см. выше). Каждый из переходов несовместен с каким-либо другим. Определим вероятность суммы несовместных событий, т. е. $P_k(t + \Delta t)$, которая будет равняться сумме вероятностей отдельных событий (по теореме сложения вероятностей несовместных событий):

$$P_k(t + \Delta t) = p_{k1}(t + \Delta t) + p_{k2}(t + \Delta t) + p_{k3}(t + \Delta t) + o(\Delta t), \quad k \geq 1. \quad (2.1)$$

В равенстве (2.1) учтена погрешность вероятностей перехода за время Δt символом $o(\Delta t)$. Условие $k \geq 1$ учитывает то, что если система находится в нулевом состоянии, т. е. когда число членов популяции равно нулю, то гибели происходить не может, что было отмечено в определении процесса размножения и гибели. С учетом вероятностей $p_{ki}(t + \Delta t)$, $i = 1, 2, 3$ и вероятностей перехода за время Δt (см. выше) представим равенство (2.1) в виде

$$P_k(t + \Delta t) = P_k(t)[1 - \lambda_k \Delta t - \mu_k \Delta t + o(\Delta t)] + P_{k-1}(t)[\lambda_{k-1} \Delta t + o(\Delta t)] + P_{k+1}(t)[\mu_{k+1} \Delta t + o(\Delta t)], \quad k \geq 1. \quad (2.2)$$

В равенстве (2.2) раскроем скобки, суммируем малые порядка выше, чем Δt и сгруппируем слагаемые в следующем виде:

$$P_k(t + \Delta t) = P_k(t) - (\lambda_k + \mu_k)P_k(t)\Delta t + P_{k-1}(t)\lambda_{k-1}\Delta t + P_{k+1}(t)\mu_{k+1}\Delta t + o(\Delta t), \quad k \geq 1. \quad (2.3)$$

Выражение (2.3) — это бесконечная система равенств для $k \geq 1$. Рассмотрим равенство (2.3) для случая $k = 0$. По определению процесса размножения и гибели состояние E_0 соответствует отсутствию членов популяции в системе. Рассмотрим возможности достижения состояния E_0 за время Δt , их только два:

1. За время Δt не происходит никаких изменений. Это условие отразим в виде:

$$\begin{array}{c} \Delta t \\ E_0 \longrightarrow E_0. \end{array}$$

2. За время Δt происходит переход в E_0 из E_1 . Это условие отразим в виде:

$$\begin{array}{c} \Delta t \\ E_1 \longrightarrow E_0. \end{array}$$

Определим вероятности переходов в состояние E_0 . Если формально осуществляется переход из E_0 в E_0 , то вероятность $P_{0,0}(\Delta t)$ такого перехода выразим как дополнение к вероятности противоположного события, а именно к рождению одного члена популяции:

$$P_{0,0}(\Delta t) = 1 - \lambda_0 \Delta t + o(\Delta t). \quad (2.4)$$

При переходе из состояния E_1 в E_0 происходит гибель одного члена популяции. Вероятность $P_{1,0}(\Delta t)$ такого перехода равна

$$P_{1,0}(\Delta t) = \mu_1 \Delta t + o(\Delta t). \quad (2.5)$$

Для момента времени $t + \Delta t$ определим вероятность состояния сначала для каждого из возможных переходов:

$$p_{0,1}(t + \Delta t) = P_0(t) \cdot P_{0,0}(\Delta t), \quad p_{0,2}(t + \Delta t) = P_1(t) \cdot P_{1,0}(\Delta t). \quad (2.6)$$

Искомая вероятность $P_0(t + \Delta t)$ определится как сумма вероятностей несовместных событий:

$$P_0(t + \Delta t) = p_{0,1}(t + \Delta t) + p_{0,2}(t + \Delta t).$$

С учетом (2.4)-(2.6) получим

$$P_0(t + \Delta t) = P_0(t)[1 - \lambda_0 \Delta t + o(\Delta t)] + P_1(t)[\mu_1 \Delta t + o(\Delta t)]. \quad (2.7)$$

В (2.7) раскроем скобки и приведем подобные члены:

$$P_0(t + \Delta t) = P_0(t) - \lambda_0 P_0(t) \Delta t + \mu_1 P_1(t) \Delta t + o(\Delta t). \quad (2.8)$$

Полученное равенство (2.8) включим в систему равенств (2.3):

$$P_k(t + \Delta t) = P_k(t) - (\lambda_k + \mu_k) P_k(t) \Delta t + P_{k-1}(t) \lambda_{k-1} \Delta t + P_{k+1}(t) \mu_{k+1} \Delta t + o(\Delta t), \quad k \geq 1, \quad (2.9)$$

$$P_0(t + \Delta t) = P_0(t) - \lambda_0 P_0(t) \Delta t + \mu_1 P_1(t) \Delta t + o(\Delta t).$$

2.2.4. Вывод дифференциальных уравнений процесса размножения и гибели

Систему равенств (2.9) поделим на Δt и запишем в виде:

$$\frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = \frac{-(\lambda_k + \mu_k)P_k(t)\Delta t}{\Delta t} + \frac{P_{k-1}(t)\lambda_{k-1}\Delta t}{\Delta t} + \frac{P_{k+1}(t)\mu_{k+1}\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t},$$

$$k \geq 1,$$

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = \frac{-\lambda_0 P_0(t)\Delta t}{\Delta t} + \frac{\mu_1 P_1(t)\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t}, \quad k = 0.$$

Далее перейдем к пределу при $\Delta t \rightarrow 0$:

$$\lim_{\Delta t \rightarrow 0} \frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = -(\lambda_k + \mu_k)P_k(t) + P_{k-1}(t)\lambda_{k-1} + P_{k+1}(t)\mu_{k+1} + \lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t},$$

$$k \geq 1,$$

$$\lim_{\Delta t \rightarrow 0} \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda_0 P_0(t) + \mu_1 P_1(t) + \lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t}, \quad k = 0.$$

Левые части равенств формально представляют собой производные от

$P_k(t)$ и $P_0(t)$ по t , предел $\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$, так как $o(\Delta t)$ есть малая более

высокого порядка, чем Δt , следовательно, предел правой части существует. В итоге получаем следующую систему обыкновенных дифференциальных уравнений с постоянными коэффициентами:

$$\begin{aligned} \frac{dP_k(t)}{dt} &= -(\lambda_k + \mu_k)P_k(t) + \lambda_{k-1}P_{k-1}(t) + \mu_{k+1}P_{k+1}(t), \quad k \geq 1, \\ \frac{dP_0(t)}{dt} &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \quad k = 0. \end{aligned} \tag{2.10}$$

Дифференциальные уравнения (2.10) описывают динамику вероятностного процесса размножения и гибели. Эти уравнения называются уравнениями А. Н. Колмогорова [8]. Для решения дифференциальных уравнений (2.10) необходимо задать начальные условия, т. е. задать $P_k(0)$ при $k = 0, 1, 2, \dots$. В соответствии с постулатами процесса размножения и гибели в системе возможно как рождение, так и гибель, т. е. состояние системы может изменяться только на соседнее состояние. Следовательно, в начальный момент времени система может находиться в произвольном состоянии E_i . Тогда вероятность этого

состояния в начальный момент времени будет равна единице, а вероятность других состояний равна нулю [15]. Поэтому если состоянием системы является E_i , то начальными условиями для решения дифференциальных уравнений (2.10) будут

$$P_i(0) = 1, \quad P_k(0) = 0, \quad k \neq i. \quad (2.11)$$

В случае, когда начало процесса определяется состоянием E_0 , то начальные условия решения уравнений (2.10) имеют вид

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k \geq 1. \quad (2.12)$$

Отметим свойства решений (2.10) с начальными условиями (2.11) или (2.12). Для произвольных заданных коэффициентов $\lambda_k \geq 0$, $\mu_k \geq 0$ всегда существует положительное решение $\{P_k(t)\}$ системы (2.10) с начальными условиями (2.11) или (2.12), такое, что $\sum_{k=0}^{\infty} P_k(t) \leq 1$. Если коэффициенты ограничены (или возрастают достаточно медленно), то это решение единственно и удовлетворяет условию регулярности [15, с.470]:

$$\sum_{k=0}^{\infty} P_k(t) = 1. \quad (2.13)$$

Условие регулярности называется еще нормировочным условием.

Запишем систему (2.10) в виде

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \\ \frac{dP_1(t)}{dt} &= \lambda_0 P_0(t) - (\lambda_1 + \mu_1) P_1(t) + \mu_2 P_2(t), \\ \frac{dP_2(t)}{dt} &= \lambda_1 P_1(t) - (\lambda_2 + \mu_2) P_2(t) + \mu_3 P_3(t), \\ &\dots\dots\dots, \\ \frac{dP_i(t)}{dt} &= \lambda_{i-1} P_{i-1}(t) - (\lambda_i + \mu_i) P_i(t) + \mu_{i+1} P_{i+1}(t), \\ &\dots\dots\dots \end{aligned} \quad (2.14)$$

В свою очередь систему (2.14), а, следовательно, и (2.10) можно представить в матричной форме

$$\frac{dP(t)}{dt} = \Lambda P(t), \quad (2.15)$$

где $P(t)$ — бесконечно мерный вектор вероятности состояний, а

$$\Lambda = \begin{bmatrix} -\lambda_0 & \mu_1 & 0 & \dots & 0 & \dots \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 & 0 & \dots & 0 & \dots \\ 0 & \lambda_1 & -(\lambda_2 + \mu_2) & \mu_3 & 0 & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \lambda_{i-1} & -(\lambda_i + \mu_i) & \mu_{i+1} & 0 & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}. \quad (2.16)$$

Матрица Λ вида (2.16) называется инфинитезимальной производящей матрицей для общего однородного процесса размножения и гибели [4]. Матрица Λ называется также матрицей переходных интенсивностей. В матрице Λ за исключением главной и соседних с ней снизу и сверху диагоналей все элементы равны нулю. Между матрицей переходных интенсивностей и размеченным графом состояний системы (рис. 2.2) существует однозначное соответствие.

В практических случаях бесконечная система уравнений (2.10) должна быть заменена конечной системой, для которой процесс размножения и гибели представляется размеченным графом состояний, показанным на рис. 2.1, если в нем указать интенсивности переходов. Конечная система дифференциальных уравнений (также называемая системой уравнений Колмогорова) имеет вид

$$\begin{aligned}
 \frac{dP_0(t)}{dt} &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \\
 \frac{dP_1(t)}{dt} &= \lambda_0 P_0(t) - (\lambda_1 + \mu_1) P_1(t) + \mu_2 P_2(t), \\
 \frac{dP_2(t)}{dt} &= \lambda_1 P_1(t) - (\lambda_2 + \mu_2) P_2(t) + \mu_3 P_3(t), \\
 &\dots\dots\dots, \\
 \frac{dP_i(t)}{dt} &= \lambda_{i-1} P_{i-1}(t) - (\lambda_i + \mu_i) P_i(t) + \mu_{i+1} P_{i+1}(t), \\
 &\dots\dots\dots, \\
 \frac{dP_n(t)}{dt} &= \lambda_{n-1} P_{n-1}(t) - \mu_n P_n(t).
 \end{aligned} \tag{2.17}$$

Соответственно, для системы уравнений (2.17) в нормировочном условии изменится верхний предел суммирования:

$$\sum_{k=0}^n P_k(t) = 1. \tag{2.18}$$

Нормировочное условие (2.18) можно использовать вместо одного (любого) из дифференциальных уравнений (2.17) [8, с.133]. Начальные условия для решения системы (2.17) составляются аналогичным образом, только с учетом того, что число вероятностей состояния для начального момента времени конечно, а именно

$$P_i(0) = 1, \quad P_k(0) = 0, \quad k \neq i, \quad i, k = \overline{1, n}. \tag{2.19}$$

2.2.5. Мнемоническое правило составления уравнений Колмогорова

Дифференциальные уравнения Колмогорова 2.10, 2.17 могут быть составлены по размеченному графу состояний системы (см. рис. 2.1, 2.2). Для этого вводится понятие "поток вероятности". *Потоком вероятности* называется такой поток, который переводит систему из одного состояния в другое соседнее и определяется как произведение вероятности $P_i(t)$

состояния E_i (или по другому обозначению S_i), из которого исходит стрелка, на интенсивность потока событий (λ_i или μ_i), переводящего систему по этой стрелке [8, с.135]. Мнемоническое правило заключается в следующем [8, с.135]: *производная вероятности любого состояния равна сумме потоков вероятности, переводящих систему в это состояние, минус сумма всех потоков вероятности, выводящих систему из этого состояния.*

Для удобства применения мнемонического правила на размеченном графе состояний системы можно рекомендовать вокруг какого-либо состояния проводить воображаемую окружность. Тогда стрелки интенсивностей, входящие в окружность будут определять положительный поток вероятности, а стрелки, выходящие из окружности, будут определять отрицательный поток вероятности. Это правило поясняется на рис. 2.4, на основе которого формируется дифференциальное уравнение относительно k -го состояния, т. е. E_k .

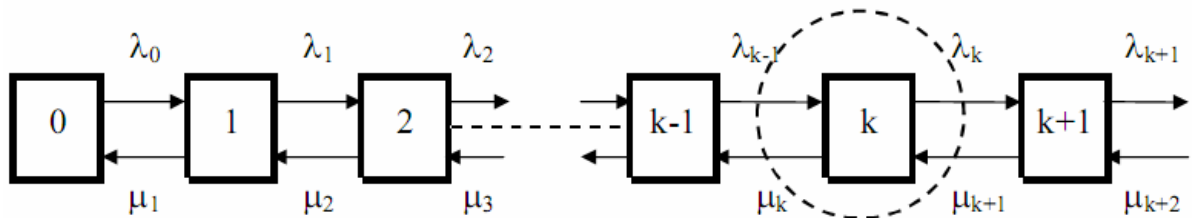


Рис. 2.4. К определению дифференциального уравнения Колмогорова по мнемоническому правилу

Для выделенного состояния на рис. 2.4 определим положительные и отрицательные потоки вероятности: положительные потоки вероятности — это $P_{k-1}\lambda_{k-1}$, $P_{k+1}\mu_{k+1}$, отрицательные потоки вероятности — это $P_k\lambda_k$, $P_k\mu_k$. Дифференциальное уравнение относительно $P_k(t)$ будет в правой части состоять из суммы положительных и отрицательных потоков вероятностей:

$$\frac{dP_k(t)}{dt} = \lambda_{k-1}P_{k-1}(t) - (\lambda_k + \mu_k)P_k(t) + \mu_{k+1}P_{k+1}(t).$$

2.2.6. Пример решения дифференциальных уравнений Колмогорова

Система дифференциальных уравнений Колмогорова в общем случае может быть решена численными методами. Для примера рассмотрим решение следующей конечной системы дифференциальных уравнений Колмогорова:

$$\begin{aligned}\frac{dP_0(t)}{dt} &= -0.5P_0(t) + 0.8P_1(t), \\ \frac{dP_1(t)}{dt} &= 0.5P_0(t) - 1.1P_1(t) + 1.2P_2(t), \\ \frac{dP_2(t)}{dt} &= 0.3P_1(t) - 1.7P_2(t) + 0.6P_3(t), \\ \frac{dP_3(t)}{dt} &= 0.5P_2(t) - 0.9P_3(t) + 0.55P_4(t), \\ \frac{dP_4(t)}{dt} &= 0.3P_3(t) - 0.55P_4(t).\end{aligned}\tag{2.20}$$

Система уравнений (2.20) — это система однородных обыкновенных дифференциальных уравнений 5-го порядка. Для ее решения зададим следующие начальные условия:

$$P_0(0) = P_1(0) = P_3(0) = P_4(0) = 0, \quad P_2(0) = 1.\tag{2.21}$$

Размеченный граф состояний для системы (2.20) будет иметь вид, показанный на рис. 2.5, где около стрелок приведены численные значения интенсивностей.

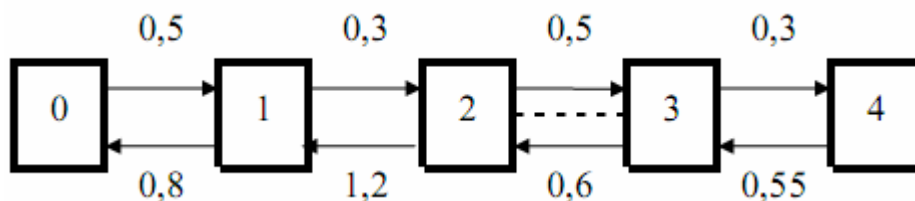


Рис. 2.5. Размеченный граф состояний системы 5-го порядка

В соответствии с системой уравнений (2.20) и графом состояний Рис. 2.5 значения интенсивностей следующие:

$$\begin{aligned}\lambda_0 &= 0,5; & \lambda_1 &= 0,3; & \lambda_2 &= 0,5; & \lambda_3 &= 0,3; \\ \mu_1 &= 0,8; & \mu_2 &= 1,2; & \mu_3 &= 0,6; & \mu_4 &= 0,55.\end{aligned}$$

Ниже приводится программа численного решения дифференциальных уравнений (2.20) в системе MATLAB версий 6.x и выше:

```
clc, close
global L
L = [-0.5 0.8 0 0 0;
     0.5 -1.1 1.2 0 0;
     0 0.3 -1.7 0.6 0;
     0 0 0.5 -0.9 0.55;
     0 0 0 0.3 -0.55];
P0 = [0;0;1;0;0];
T = [0, 20];
[t,P] = ode45(@atest11,T,P0);
plot(t,P(:,1),'k.','linewidth',2), hold on,
plot(t,P(:,2),'k:','linewidth',2.5),
plot(t,P(:,3),'k','linewidth',2),
plot(t,P(:,4),'k--','linewidth',2),
plot(t,P(:,5),'k-.','linewidth',2),
xlabel('\bf Time'),
ylabel('\bf Probability'),
legend('P_0','P_1','P_2','P_3','P_4'),
text(2.1,0.85,'P_0(0)=P_1(0)=0; P_2(0)=1; P_3(0)=P_4(0)=0'),
set(gcf,'color','w')
```


Отдельно создаваемый М-файл **atest11** описания правых частей дифференциальных уравнений (2.20) имеет вид

```
function f = atest11(t,P);  
global L  
f = L*P;
```

Результаты численного решения системы (2.20) представлены в графическом виде на рис. 2.6, где различным начертанием изображены вероятности в функции от времени. Для рассматриваемой системы существует стационарный режим, когда с течением времени вероятности остаются без изменений.

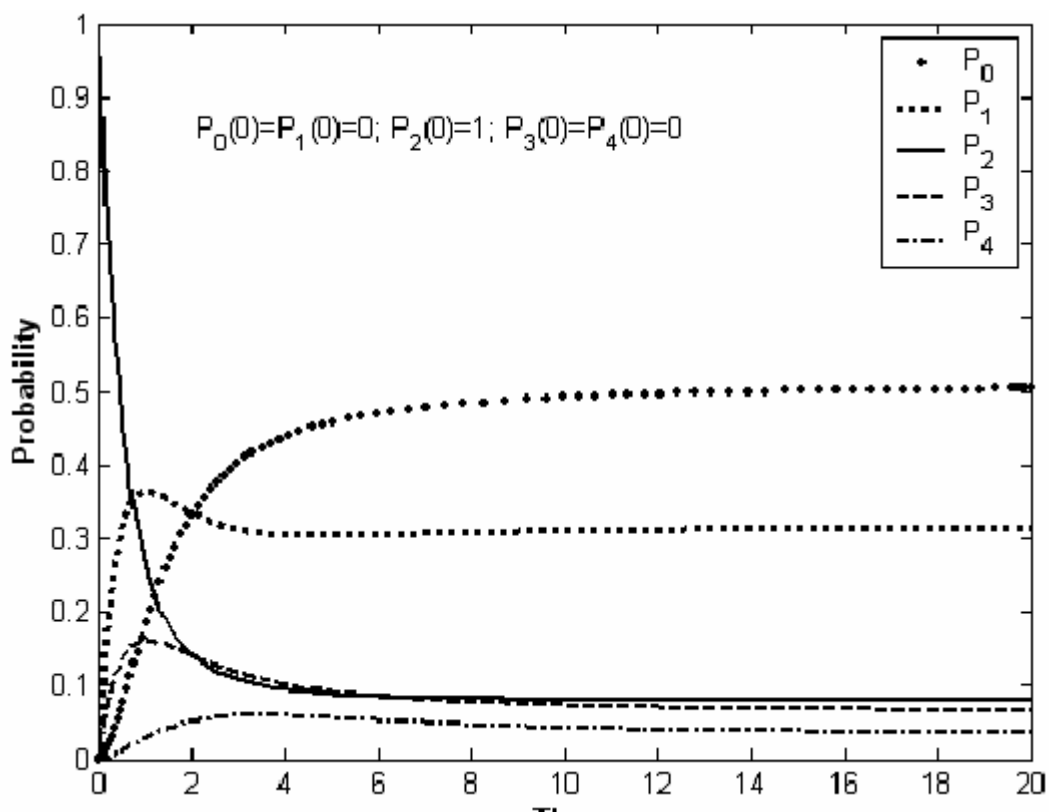


Рис. 2.6. Зависимости вероятностей от времени в системе 5-го порядка

В соответствии с нормировочным условием в каждый момент времени сумма всех пяти вероятностей должна равняться единице. Для проверки этого условия приведем некоторые результаты численного

интегрирования (с 13 по 22 шаг интегрирования), выполненные в системе MATLAB, и сведем эти результаты в табл. 2.1.

Таблица 2.1

P_0	P_1	P_2	P_3	P_4	Сумма вероятностей
0,0000	0,0062	0,9912	0,0026	0,0000	1,0000
0,0001	0,0123	0,9825	0,0051	0,0000	1,0000
0,0001	0,0184	0,9738	0,0077	0,0000	1,0000
0,0002	0,0243	0,9652	0,0102	0,0000	0,9999
0,0003	0,0302	0,9568	0,0126	0,0000	0,9999
0,0012	0,0583	0,9158	0,0244	0,0002	0,9999
0,0027	0,0845	0,8769	0,0355	0,0004	1,0000
0,0047	0,1087	0,8402	0,0457	0,0007	1,0000
0,0071	0,1311	0,8053	0,0553	0,0011	1,0000
0,0119	0,1639	0,7529	0,0694	0,0019	1,0000

Как видно из табл. 2.1, для каждого момента времени (или шага интегрирования) сумма вероятностей по каждой из строк таблицы практически равна единице. Приведенное решение системы дифференциальных уравнений (2.20) относительно вероятностей состояний удовлетворяет нормировочному условию (2.18), согласно которому сумма всех вероятностей равна единице. По графику Рис. 2.6 можно визуально оценить, что для любого момента времени сумма ординат всех пяти графиков приблизительно равна единице.

2.3. Процесс чистого размножения

Математическая модель процесса чистого размножения легко получается из модели процесса размножения и гибели, если положить, что интенсивность гибели в системе (в объеме популяции) равна нулю, т. е. $\mu_k = 0, \forall k$ (при всех k). Как и ранее будем считать, что $P_k(t)$ является вероятностью того, что система в момент времени t находится в состоянии E_k . Тогда функции $P_k(t)$ должны удовлетворять следующей системе дифференциальных уравнений:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda_0 P_0(t), \\ \frac{dP_1(t)}{dt} &= \lambda_0 P_0(t) - \lambda_1 P_1(t), \\ \frac{dP_2(t)}{dt} &= \lambda_1 P_1(t) - \lambda_2 P_2(t), \\ &\dots\dots\dots, \\ \frac{dP_i(t)}{dt} &= \lambda_{i-1} P_{i-1}(t) - \lambda_i P_i(t), \\ &\dots\dots\dots \end{aligned} \tag{2.22}$$

Система (2.22) — это основная система дифференциальных уравнений, описывающая процесс чистого размножения. Для решения системы (2.22) зададим начальные условия, определяющие начальное состояние системы:

$$P_i(0) = 1, \quad P_k(0) = 0, \quad k \neq i. \tag{2.23}$$

Задание начальных условий вида (2.23) означает, что система будет выходить из произвольного начального состояния E_i . Решение системы (2.22) в квадратурах можно получить при последовательном интегрировании дифференциальных уравнений, начиная с первого — относительно $P_0(t)$. Первое уравнение — однородное дифференциальное уравнение первого порядка, последующие — неоднородные дифференциальные уравнения также первого порядка.

Систему дифференциальных уравнений (2.22) можно получить также на основе *постулатов чистого размножения*:

1. Непосредственный переход из состояния E_i возможен только в состояние E_{i+1} .
2. Если в момент времени t система находится в состоянии E_k , то (условная) вероятность одного скачка в последующем коротком интервале времени между t и $t + \Delta t$ равна $\lambda_k \Delta t + o(\Delta t)$, тогда как вероятность (условная) более чем одного скачка в этом интервале есть $o(\Delta t)$ [15].

В силу последовательной структуры дифференциальных уравнений (2.22) их решение в определенной степени зависит от начальных условий: для общего вида начальных условий (2.23) все состояния по номеру меньшие E_i будут равны нулю на всем протяжении интервала интегрирования системы (2.22). Рассмотрим это замечание на примере моделирования системы чистого размножения, которая описывается системой дифференциальных уравнений 5-го порядка.

Пусть задана следующая система:

$$\begin{aligned}\frac{dP_0(t)}{dt} &= -0.5P_0(t), \\ \frac{dP_1(t)}{dt} &= 0.5P_0(t) - 0.3P_1(t), \\ \frac{dP_2(t)}{dt} &= 0.3P_1(t) - 0.7P_2(t), \\ \frac{dP_3(t)}{dt} &= 0.7P_2(t) - 0.9P_3(t), \\ \frac{dP_4(t)}{dt} &= 0.9P_3(t) - 0.55P_4(t).\end{aligned}\tag{2.24}$$

Сначала для решения (2.24) зададим начальные условия вида

$$P_0(0) = 1, \quad P_1(0) = P_2(0) = P_3(0) = P_4(0) = 0.\tag{2.25}$$

В MATLAB составим программу решения системы (2.24) с начальными условиями (2.25):

```
clear, clc
global L1
L1=[-0.5, 0, 0, 0, 0;
    0.5, -0.3, 0, 0, 0;
    0, 0.3, -0.7, 0, 0;
    0, 0, 0.7, -0.9, 0;
    0, 0, 0, 0.9, -0.55];
T1 =[0, 20];
P0 = [1;0;0;0;0];
[t,P1] = ode45(@atest22,T1,P0);
plot(t,P1(:,1),'k','linewidth',1.5),
hold on,
plot(t,P1(:,2),'k:', 'linewidth',2.5),
plot(t,P1(:,3),'k.', 'linewidth',2),
plot(t,P1(:,4),'k--', 'linewidth',2),
plot(t,P1(:,5),'k-.', 'linewidth',2),

grid off,
xlabel('\bf Time'),ylabel('\bf Probability'),
legend('P_0','P_1','P_2','P_3','P_4'),
text(2.1,0.85,'P_0(0)=1; P_1(0)=P_2(0)=P_3(0)=P_4(0)=0'),
set(gcf, 'color', 'w')
```

Описание правых частей дифференциальных уравнений выполнено с помощью M-файла с именем atest22:

```
function f = atest22(t,P1);
```

global L1

f = L1*P1;

Графические результаты решения системы (2.24) с начальными условиями (2.25) представлены на рис. 2.7.

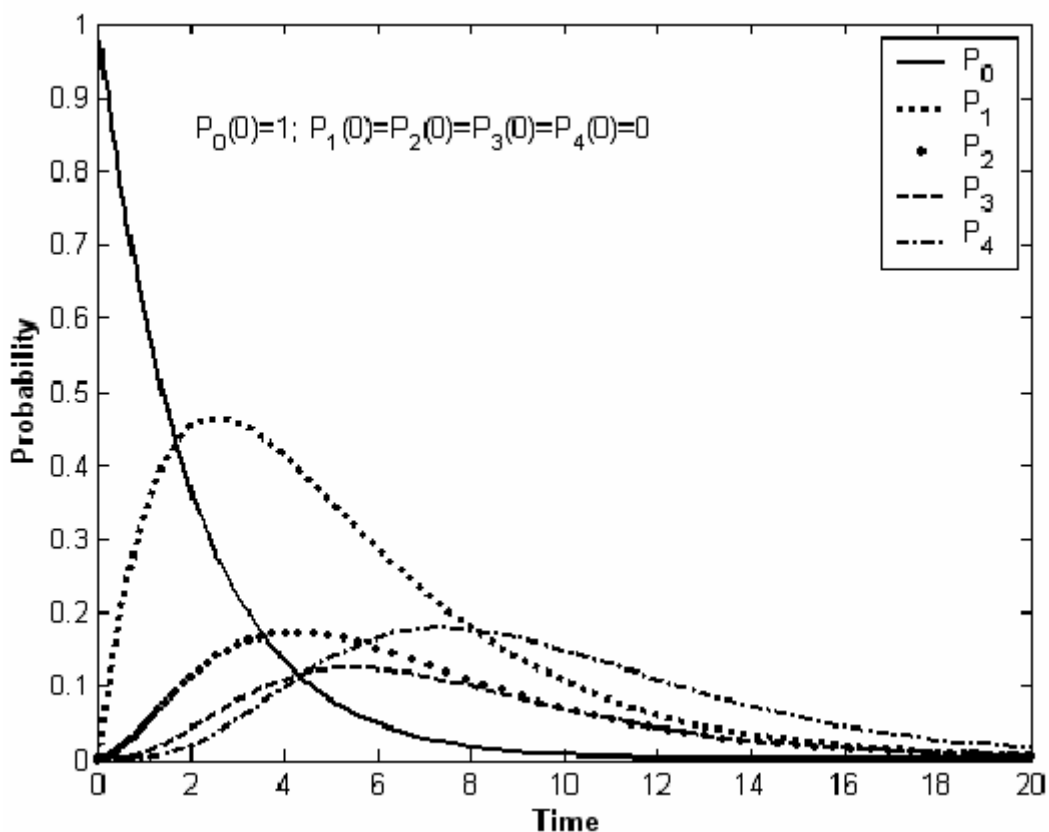


Рис. 2.7. Изменение вероятностей состояний процесса чистого размножения с начальными условиями:

$$P_0(0) = 1, \quad P_1(0) = P_2(0) = P_3(0) = P_4(0) = 0$$

Для той же системы (2.24) зададим другие начальные условия:

$$P_0(0) = P_2(0) = P_3(0) = P_4(0) = 0, \quad P_1(0) = 1. \quad (2.26)$$

Численное решение системы (2.24) с начальными условиями (2.26) можно также определить по приведенной MATLAB-программе. Графические результаты решения системы (2.24) с начальными условиями (2.26) представлены на рис. 2.8, где видно, что вероятность состояния $P_0(t) = 0$ для всего заданного интервала интегрирования. Это согласуется и со

структурой заданных дифференциальных уравнений (2.24) и с постулатами процесса чистого размножения, когда переход из одного состояния возможен только в следующее по номеру больше исходного.

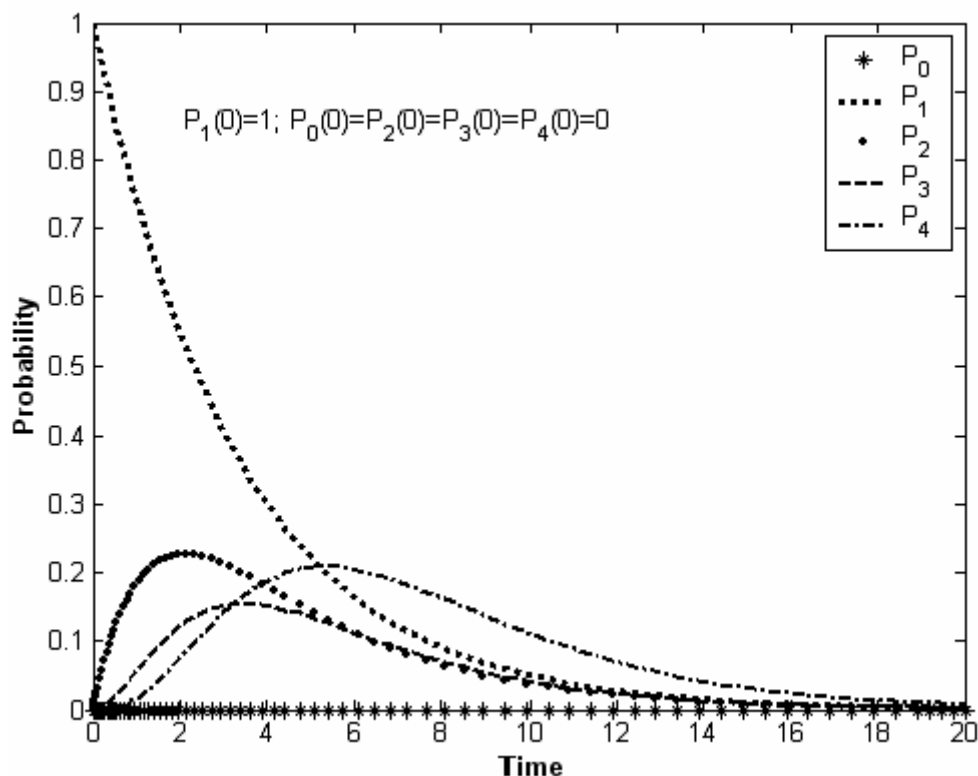


Рис. 2.8. Изменение вероятностей состояний процесса чистого размножения с начальными условиями:

$$P_0(0) = P_2(0) = P_3(0) = P_4(0) = 0, \quad P_1(0) = 1.$$

Как видно из рисунков 2.7, 2.8 вероятность того, что с течением времени в системе будет 4 члена популяции практически равна единице. В соответствии с последними начальными условиями (2.26) можно утверждать, что если сначала в системе было три члена популяции (вероятность этого равна единице), то с течением времени в системе становится четыре члена популяции, т.к. вероятность этого события стремится к единице. Вероятности остальных состояний равны нулю (или становятся равными нулю).

Если для уравнений (2.22) процесса чистого размножения принять одинаковые интенсивности $\lambda_k = \lambda = const$, то получим следующую систему дифференциальных уравнений:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda P_0(t), & k=0, \\ \frac{dP_k(t)}{dt} &= -\lambda P_k(t) + \lambda P_{k-1}(t), & k \geq 1. \end{aligned} \quad (2.27)$$

Решение системы (2.27) может быть представлено в квадратурах при последовательном интегрировании, начиная с первого. Последовательное решение до k дает распределение Пуассона:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k \geq 0, \quad t \geq 0. \quad (2.28)$$

Для получения распределения Пуассона (2.28) из системы (2.27) следует допустить, что в момент времени 0 система выходит из начального состояния E_0 . Это означает, что начальные условия при решении системы дифференциальных уравнений (2.27) должны быть следующими:

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k \geq 1. \quad (2.29)$$

2.4. Процесс чистой гибели

Процессом чистой гибели называется такой процесс размножения и гибели, у которого интенсивности всех потоков размножения равны нулю, т. е. когда

$$\lambda_k = 0, \quad \mu_k \neq 0, \quad \forall k. \quad (2.30)$$

Размеченный граф процесса чистой гибели с конечным числом состояний показан на рис. 2.9.

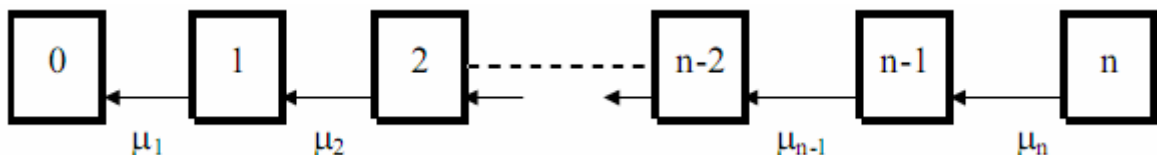


Рис. 2.9. Размеченный граф состояний процесса чистой гибели

Система уравнений, описывающих процесс чистой гибели, получается из общей системы конечного числа уравнений Колмогорова, если для нее принять условия (2.30), т. е.

$$\begin{aligned}\frac{dP_0(t)}{dt} &= \mu_1 P_1(t), \\ \frac{dP_1(t)}{dt} &= -\mu_1 P_1(t) + \mu_2 P_2(t), \\ \frac{dP_2(t)}{dt} &= -\mu_2 P_2(t) + \mu_3 P_3(t), \\ &\dots\dots\dots, \\ \frac{dP_i(t)}{dt} &= -\mu_i P_i(t) + \mu_{i+1} P_{i+1}(t), \\ &\dots\dots\dots, \\ \frac{dP_n(t)}{dt} &= -\mu_n P_n(t).\end{aligned}\tag{2.31}$$

При составлении дифференциальных уравнений можно пользоваться мнемоническим правилом: *производная вероятности любого состояния равна сумме потоков вероятности, переводящих систему в это состояние, минус сумма всех потоков вероятности, выводящих систему из этого состояния*. В данном случае под потоком вероятности понимается произведение вероятности состояния на интенсивность гибели, причем поток вероятности положителен, если стрелка интенсивности входит в заданное состояние и отрицателен в противном случае.

Решение системы (2.31) производится при задании начальных условий вида

$$P_i(0) = 1, \quad P_k(0) = 0, \quad k \neq i, \quad k = \overline{1, n}.\tag{2.32}$$

Естественными начальными условиями для решения дифференциальных уравнений (2.31) являются такие, когда задается начальный объем популяции, в котором начинается происходить гибель:

$$P_n(0) = 1, \quad P_k(0) = 0, \quad k = \overline{1, n-1}. \quad (2.33)$$

Рассмотрим пример процесса чистой гибели, который описывается следующей системой дифференциальных уравнений:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= 2P_1(t), \\ \frac{dP_1(t)}{dt} &= -2P_1(t) + 3,33P_2(t), \\ \frac{dP_2(t)}{dt} &= -3,33P_2(t) + 7,9P_3(t), \\ \frac{dP_3(t)}{dt} &= -7,9P_3(t) + 6,55P_4(t), \\ \frac{dP_4(t)}{dt} &= -6,55P_4(t), \end{aligned} \quad (2.34)$$

где $\mu_1 = 2$; $\mu_2 = 3,33$; $\mu_3 = 7,9$; $\mu_4 = 6,55$.

Для решения системы (2.34) зададим начальные условия:

$$P_4(0) = 1, \quad P_0(0) = P_1(0) = P_2(0) = P_3(0) = 0. \quad (2.35)$$

Программный код решения примера:

```
clear,clc, close
global L2
L2 = [0, 2, 0, 0, 0;
0, -2, 3.33, 0, 0;
0, 0, -3.33, 7.9, 0;
0, 0, 0, -7.9, 6.55;
0, 0, 0, 0, -6.55];
T2 = [0, 10];
P0 = [0;0;0;0;1];
[t,P2] = ode45(@atest33,T2,P0);
```

```
plot(t,P2(:,1),'k--','linewidth',2), hold on,  
plot(t,P2(:,2),'k:', 'linewidth',2.5),  
plot(t,P2(:,3),'k.', 'linewidth',2),  
plot(t,P2(:,4),'k-', 'linewidth',2),  
plot(t,P2(:,5),'k', 'linewidth',2),  
grid off,  
xlabel('\bf Time'),  
ylabel('\bf Probability'),  
legend('P_0','P_1','P_2','P_3','P_4',4),  
text(2.1,0.85,'P_0(0)=P_1(0)=P_2(0)=P_3(0)=0; P_4(0)=1'),  
axis([0, 10, 0, 1.2]),  
set(gcf,'color','w')
```

M-файл описания правых частей дифференциальных уравнений:

```
function f = atest33(t,P2);  
global L2  
f = L2*P2;
```

Графическое представление решение системы (2.34) с начальными условиями (2.35) показано на рис. 2.9.

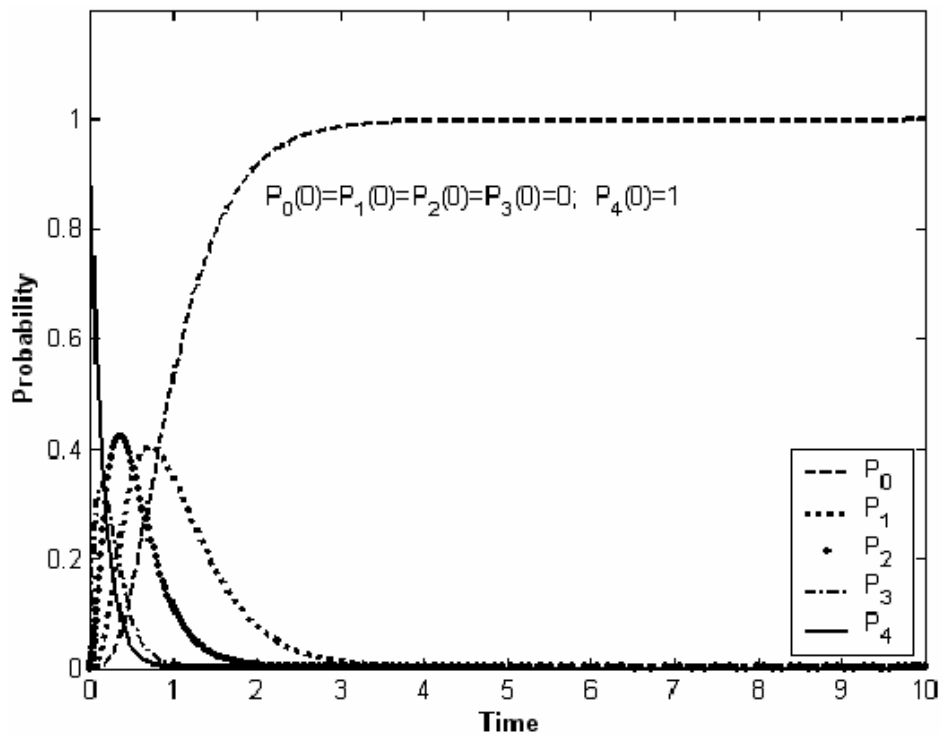


Рис. 2.9. Изменение вероятностей состояний процесса чистой гибели

Как видно из рис. 2.9 гибель членов популяции в системе подтверждается тем, что вероятность $P_0(t)$ отсутствия членов популяции стремится к единице, а остальные вероятности $P_1(t)$, $P_2(t)$, $P_3(t)$, $P_4(t)$ — вероятности 1-го, 2-х, 3-х, 4-х членов популяции стремятся к нулю.

Если процесс гибели начинается в момент, когда объем популяции равен N и характеризуется одинаковой и постоянной интенсивностью $\mu_k = \mu = \text{const}$, то система уравнений (2.31) процесса чистой гибели будет иметь вид

$$\begin{aligned} \frac{dP_0(t)}{dt} &= \mu P_1(t), & k=0; \\ \frac{dP_k(t)}{dt} &= -\mu P_k(t) + \mu P_{k+1}(t), & 0 < k < N; \\ \frac{dP_N(t)}{dt} &= -\mu P_N(t), & k=N. \end{aligned} \quad (2.36)$$

Для системы дифференциальных уравнений, описывающих процесс чистой гибели, начиная с объема популяции N , начальные условия будут вида

$$P_N(0) = 1, \quad P_k(0) = 0, \quad k = \overline{0, N-1}. \quad (2.37)$$

Система уравнений (2.36) с начальными условиями (2.37) легко интегрируется последовательно уравнение за уравнением, начиная с последнего уравнения. Решение последнего уравнения системы (2.36) имеет вид

$$P_N(t) = e^{-\mu t}. \quad (2.38)$$

Остальные решения находятся при подстановке в уравнение предыдущего решения. Каждое из уравнений является линейным неоднородным уравнением первого порядка. В результате решения каждого из уравнений в итоге получим:

$$\begin{aligned} P_k(t) &= \frac{(\mu t)^{N-k}}{(N-k)!} e^{-\mu t}, \quad 1 \leq k \leq N; \\ \frac{dP_0(t)}{dt} &= \frac{\mu(\mu t)^{N-1}}{(N-1)!} e^{-\mu t}, \quad k = 0. \end{aligned} \quad (2.39)$$

Вероятность нулевого состояния $P_0(t)$ должна определяться из решения дифференциального уравнения. Однако она может быть найдена из нормировочного условия:

$$\sum_{k=0}^N P_k(t) = 1, \Rightarrow P_0(t) + \sum_{k=1}^N P_k(t) = 1, \Rightarrow P_0(t) = 1 - \sum_{k=1}^N P_k(t). \quad (2.40)$$

2.5. Стационарный режим процесса размножения и гибели

Рассмотренные в примерах процессы размножения и гибели характеризуются временными зависимостями вероятностей состояний. В приводимых диаграммах (Рис. 2.6-2.9) вероятности состояний с течением времени становятся практически постоянными, не зависящими от времени. В этом смысле установившийся режим относительно

вероятностей состояний называется стационарным режимом. Говорят, что система обладает стационарным режимом, если существуют пределы

$$\lim_{t \rightarrow \infty} P_k(t) = p_k = \text{const}, \quad \forall k, \quad (2.41)$$

где p_k — предельные или финальные вероятности. Предельная вероятность p_k — это вероятность того, что в произвольный момент достаточно отдаленного будущего система будет находиться в состоянии E_k . Пределы (2.41) не зависят от начальных условий, при которых решаются дифференциальные уравнения, описывающие случайный процесс (в данном случае процесс размножения и гибели). В дальнейшем для систем массового обслуживания p_k будет означать вероятность того, что в системе находится k требований.

В стационарном режиме должно выполняться нормирующее условие относительно предельных вероятностей:

$$\sum_{k=0}^{\infty} p_k = 1. \quad (2.42)$$

Существование пределов вида (2.41) в общем случае основано на эргодической теореме Маркова [7]. Отметим, что множество состояний всей системы называется эргодическим (от греч. — работа и путь), если из любого состояния E_i можно перейти в любое другое состояние E_j [8]. Процесс размножения и гибели является частным случаем марковского процесса с дискретными состояниями и непрерывным временем. Случайный процесс с дискретными состояниями и непрерывным временем называется *марковским*, если для любого момента времени t условные вероятности всех состояний системы в будущем (при $t > t_0$) зависят только от того, в каком состоянии находится система в настоящем (при $t = t_0$), но не зависят от того, когда и каким образом она пришла в это состояние [8]. Известно, что любой марковский процесс с

конечным числом состояний, обладающий эргодическим свойством, имеет стационарный режим, который обязательно наступит после достаточного времени функционирования системы [8, с. 149]. В свою очередь, чтобы марковский процесс обладал эргодическим свойством, необходимо выполнение двух условий:

1. Граф состояний системы не должен иметь ни одного состояния и ни одного подмножества состояний без выхода и без входа (множество всех состояний системы должно быть эргодическим).
2. Все потоки событий, переводящие систему из состояния в состояние, должны быть простейшими с постоянными интенсивностями (марковский процесс должен быть однородным, т. е. с постоянными переходными интенсивностями) [8].

Системы, в которых протекает эргодический марковский случайный процесс, принято называть *простейшими эргодическими системами* [8].

Для процесса размножения и гибели найдем формальные условия существования стационарного режима. В случае существования предельных вероятностей их можно определить из уравнений Колмогорова, в которых следует положить равными нулю производные от предельных вероятностей, не зависящих от времени t :

$$\begin{aligned}
 0 &= -\lambda_0 p_0 + \mu_1 p_1, \\
 0 &= \lambda_0 p_0 - (\lambda_1 + \mu_1) p_1 + \mu_2 p_2, \\
 0 &= \lambda_1 p_1 - (\lambda_2 + \mu_2) p_2 + \mu_3 p_3, \\
 &\dots\dots\dots, \\
 0 &= \lambda_{k-1} p_{k-1} - (\lambda_k + \mu_k) p_k + \mu_{k+1} p_{k+1}, \\
 &\dots\dots\dots
 \end{aligned}
 \tag{2.43}$$

Начнем решение алгебраической системы (2.43) с первого уравнения относительно p_1 при $k = 0$:

$$p_1 = \frac{\lambda_0}{\mu_1} p_0. \quad (2.44)$$

С учетом (2.44) решим следующее уравнение относительно p_2 при $k = 1$:

$$0 = \lambda_0 p_0 - (\lambda_1 + \mu_1) p_1 + \mu_2 p_2;$$

$$0 = \lambda_0 p_0 - (\lambda_1 + \mu_1) \frac{\lambda_0}{\mu_1} p_0 + \mu_2 p_2;$$

$$0 = \lambda_0 p_0 - \frac{\lambda_1 \lambda_0}{\mu_1} p_0 - \lambda_0 p_0 + \mu_2 p_2,$$

откуда

$$p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0. \quad (2.45)$$

Далее, по индукции можно прийти к решению относительно p_k :

$$p_k = \frac{\lambda_0 \lambda_1 \dots \lambda_{k-1}}{\mu_1 \mu_2 \dots \mu_k} p_0$$

или

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}, \quad k = 0, 1, 2, \dots \quad (2.46)$$

В формуле (2.46) все вероятности p_k для стационарного режима выражаются через одну неизвестную вероятность p_0 . Используя нормировочное условие (2.42), найдем p_0 :

$$\sum_{k=0}^{\infty} p_k = 1, \quad p_0 + \sum_{k=1}^{\infty} p_k = 1, \quad p_0 + \sum_{k=1}^{\infty} p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} = 1, \quad p_0 \left(1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) = 1,$$

$$p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}. \quad (2.47)$$

Вероятность, как число, должна быть заключена в интервале $[0; 1]$. Как видно из (2.47), значение p_0 при любых λ_k и μ_k не будет превосходить 1.

Значения вероятностей p_k не будут превосходить 1 только в том случае, если выполняется условие

$$\frac{\lambda_k}{\mu_{k+1}} < 1, \quad k \geq k_0 > 0. \quad (2.48)$$

С выполнением неравенства (2.48) произведение, определяемое в (2.47), всегда будет не более 1 и, следовательно, вероятности p_k также не будут превосходить единицы.

Таким образом, условием существования стационарного режима процесса размножения и гибели является неравенство (2.48). Полученные формулы расчета предельных вероятностей (2.46), (2.47) с неравенством (2.48) являются основополагающими в анализе систем массового обслуживания, если они являются простейшими эргодическими системами.

Системы, находящиеся в стационарном режиме могут быть представлены в виде размеченного графа состояний, который ничем не будет отличаться от графа состояний, рассмотренного выше (рис. 2.2). По заданному графу состояний системы или по диаграмме интенсивностей переходов достаточно просто могут быть составлены стационарные уравнения относительно предельных вероятностей состояний. Для этого следует придерживаться следующего мнемонического правила [8, с.192]: *для простейшей схемы гибели и размножения, находящейся в стационарном режиме, потоки вероятностей между любыми двумя соседними состояниями равны.* При этом можно каждое состояние очертить, например, воображаемой окружностью и считать, сколько потоков вероятностей входит в окружность и сколько потоков вероятностей выходит из окружности. По-другому, стрелки интенсивностей, входящие в окружность формируют положительный поток вероятностей, а стрелки интенсивностей, выходящие из

воображаемой окружности, формируют отрицательный поток вероятностей. Напомним, что поток вероятности — это произведение вероятности p_i (предельной вероятности) состояния E_i , из которого исходит стрелка, на интенсивность потока событий (λ_i или μ_i), переводящего систему по этой стрелке. Например, для второго уравнения системы (2.43) можем записать следующее равенство потоков предельных вероятностей:

$$(\lambda_1 + \mu_1)p_1 = \lambda_0 p_0 + \mu_2 p_2. \quad (2.49)$$

Равенство вида (2.49) означает, что для стационарного режима в каждом состоянии входящий поток вероятностей должен быть равен исходящему потоку вероятностей.

ГЛАВА 3. ОДНОЛИНЕЙНЫЕ СИСТЕМЫ ОБСЛУЖИВАНИЯ

В данной главе будут рассматриваться в основном системы массового обслуживания (СМО), которые описываются процессами размножения и гибели. При этом процесс размножения будет соответствовать поступлению в систему требований по закону Пуассона, а гибель будет соответствовать выходу из системы обслуженных требований по экспоненциальному вероятностному закону. *Однолинейные или одноканальные системы обслуживания — это такие системы, в которых обслуживание требований осуществляется одним прибором, одним каналом, одной линией, одним узлом и т.п.* В символике обозначений таких систем обязательно присутствуют первые три поля: $M/M/1$, где первая буква (поле) M означает пуассоновский закон поступления в систему требований, вторая буква M — экспоненциальный закон обслуживания требований, цифра 1 — один канал (прибор) обслуживания. В обозначениях могут присутствовать и другие поля, которые будут раскрываться по мере изучения соответствующих систем.

3.1. Система массового обслуживания $M/M/1$. Распределение вероятностей в системе $M/M/1$

Систему $M/M/1$ принято называть классической системой массового обслуживания. В соответствии с обозначением эта система допускает бесконечную очередь, поэтому иногда ее обозначают как $M/M/1/\infty$. Система может быть описана через уравнения процесса

размножения и гибели, если установить следующие значения для коэффициентов [4]:

$$\begin{aligned}\lambda_k &= \lambda = \text{const}, & k = 0, 1, 2, \dots; \\ \mu_k &= \mu = \text{const}, & k = 1, 2, 3, \dots\end{aligned}\quad (0.1)$$

Если принять, что пространство состояний бесконечно, то размеченный граф состояний для системы М/М/1 будет иметь вид, показанный на рис. 3.1.

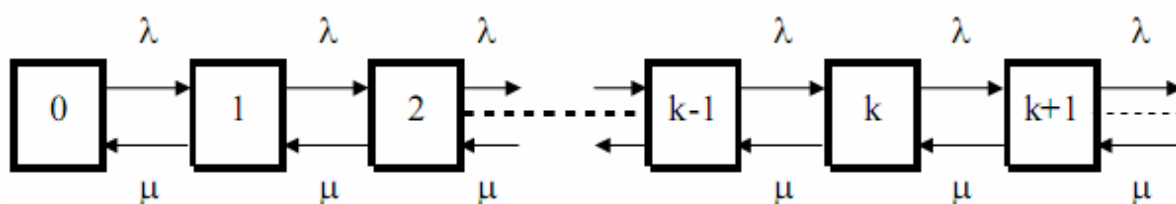


Рис. 3.1. Граф состояний для системы М/М/1

Привлечем сначала аппарат дифференциальных уравнений процесса гибели и размножения для описания процессов в системе М/М/1. Для этого запишем систему уравнений размножения и гибели (2.14) с учетом (3.1):

$$\begin{aligned}\frac{dP_0(t)}{dt} &= -\lambda P_0(t) + \mu P_1(t), \\ \frac{dP_1(t)}{dt} &= \lambda P_0(t) - (\lambda + \mu)P_1(t) + \mu P_2(t), \\ \frac{dP_2(t)}{dt} &= \lambda P_1(t) - (\lambda + \mu)P_2(t) + \mu P_3(t), \\ &\dots\dots\dots, \\ \frac{dP_k(t)}{dt} &= \lambda P_{k-1}(t) - (\lambda + \mu)P_k(t) + \mu P_{k+1}(t), \\ &\dots\dots\dots\end{aligned}\quad (0.2)$$

Условием существования стационарного режима работы системы М/М/1, а, следовательно, и установившегося решения системы (3.2) будет следующее неравенство (см. неравенство 2.48):

$$\frac{\lambda}{\mu} < 1. \quad (0.3)$$

Для стационарного режима работы системы М/М/1 используем распределения предельных вероятностей (2.46), (2.47) с учетом (3.1):

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu} = p_0 \underbrace{\frac{\lambda}{\mu} \cdot \frac{\lambda}{\mu} \cdot \dots \cdot \frac{\lambda}{\mu}}_k = p_0 \left(\frac{\lambda}{\mu} \right)^k, \quad k = 0, 1, 2, \dots \quad (0.4)$$

$$p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda}{\mu}} = \frac{1}{1 + \sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu} \right)^k} = \frac{1}{1 + \frac{\lambda}{\mu} \cdot \frac{1}{1 - \frac{\lambda}{\mu}}} = \frac{1}{1 + \frac{\lambda}{\mu} \cdot \frac{\mu}{\mu - \lambda}} = 1 - \frac{\lambda}{\mu}. \quad (0.5)$$

Так как $\lambda < \mu$, то сумма в (3.5) сходится по формуле убывающей геометрической прогрессии со знаменателем, равным λ/μ , т. е.

$$\sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu} \right)^k = \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu} \right)^2 + \left(\frac{\lambda}{\mu} \right)^3 + \dots = \frac{\lambda/\mu}{1 - \frac{\lambda}{\mu}}. \quad (0.6)$$

Введем следующее обозначение, принимая во внимание (3.3):

$$\rho = \frac{\lambda}{\mu} < 1. \quad (0.7)$$

Тогда перепишем (3.4) с учетом (3.5) и (3.7):

$$p_k = (1 - \rho) \rho^k, \quad k = 0, 1, 2, \dots \quad (0.8)$$

Вероятности, вычисляемые по формуле (3.8), определяют собой решение системы (3.2) в установившемся режиме. Вероятность p_k — это вероятность того, что в системе М/М/1 будет находиться точно k требований. Предельные вероятности p_k будем также называть стационарными вероятностями. Характерный вид распределения стационарных вероятностей в системе М/М/1 показан на рис. 3.2.

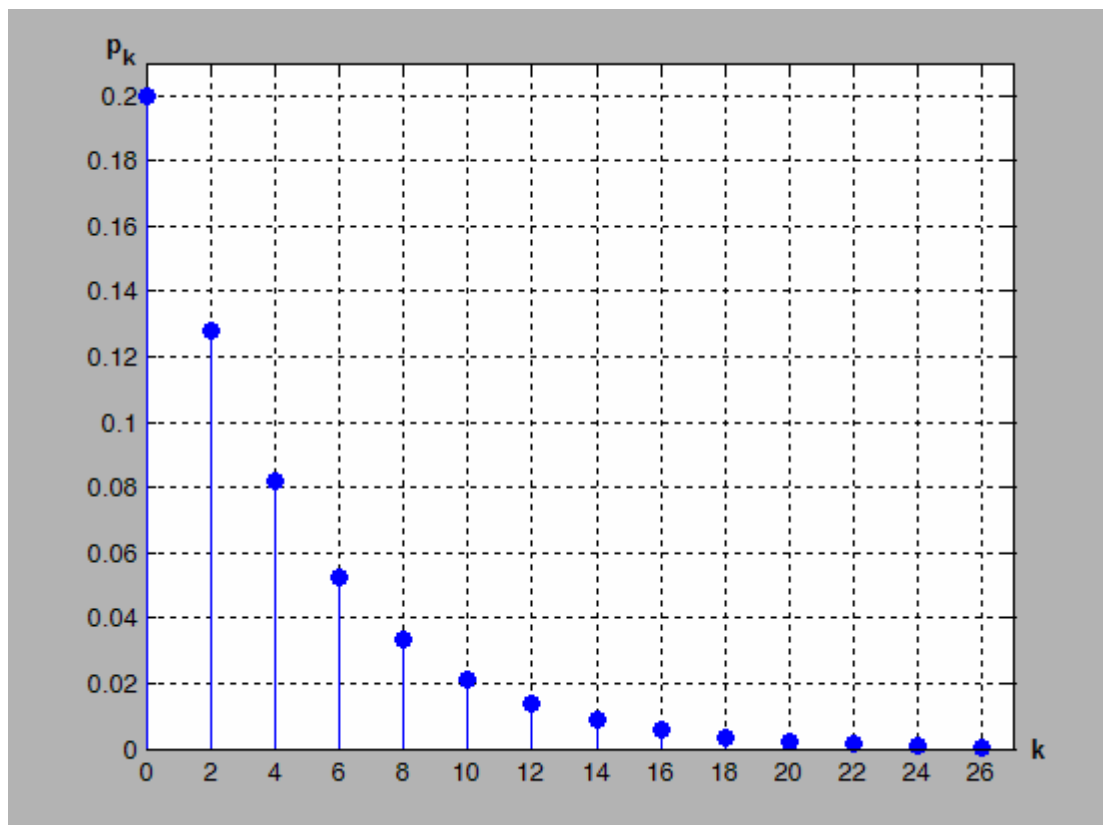


Рис. 3.2. Стационарные вероятности системы M/M/1

Стационарные вероятности для системы M/M/1 имеют дискретное распределение.

Рассмотрим характер изменения вероятностей состояния системы M/M/1/5 во времени в соответствии с графом состояний, приведенном на рис. 3.3.

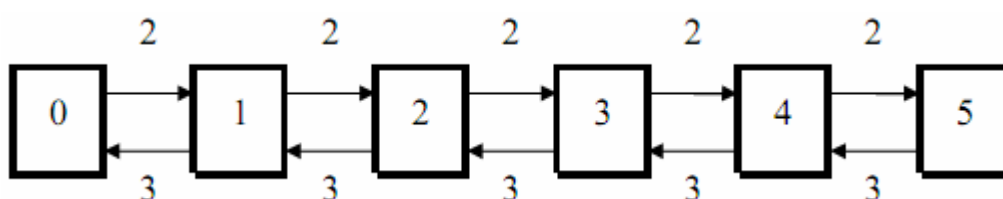


Рис. 3.3. Граф состояний для системы M/M/1/5

Система M/M/1/5 соответствует случаю, когда в системе не может быть более 5 требований. Для этой системы граф состояний конечный.

Рассмотрим пример. По заданному графу состояний рис. 3.3 составим систему дифференциальных уравнений, принимая во внимание

(3.2), или, применяя мнемоническое правило составления дифференциальных уравнений для процесса размножения и гибели:

$$\begin{aligned}\frac{dP_0(t)}{dt} &= -2P_0(t) + 2P_1(t), \\ \frac{dP_1(t)}{dt} &= 2P_0(t) - 5P_1(t) + 3P_2(t), \\ \frac{dP_2(t)}{dt} &= 2P_1(t) - 5P_2(t) + 3P_3(t), \\ \frac{dP_3(t)}{dt} &= 2P_2(t) - 5P_3(t) + 3P_4(t), \\ \frac{dP_4(t)}{dt} &= 2P_3(t) - 5P_4(t) + 3P_5(t), \\ \frac{dP_5(t)}{dt} &= 2P_4(t) - 3P_5(t).\end{aligned}\tag{0.9}$$

Для решения системы (3.9) зададим следующие начальные условия:

$$P_0(0) = 1, \quad P_1(0) = P_2(0) = P_3(0) = P_4(0) = P_5(0) = 0.\tag{0.10}$$

Программный код решения примера (в виде М-функции):

```
function MM1
clc,close
%% Параметры системы М/М1
L = 2; M = 3;
A = [-2, 3, 0, 0, 0, 0;
      2, -5, 3, 0, 0, 0;
      0, 2, -5, 3, 0, 0;
      0, 0, 2, -5, 3, 0;
      0, 0, 0, 2, -5, 3;
      0, 0, 0, 0, 2, -3];
%% Начальный вектор состояния
P0 = [1;0;0;0;0;0];
%% Интервал интегрирования
```

```
T = [0, 15];
[t,P] = ode45(@cmo11,T,P0,[],A);
%% Финальные вероятности
P(end,:)

%% Диаграмма изменения вероятностей во времени
plot(t,P(:,1),'linew',2.5),
hold on,
plot(t,P(:,2),'x','linew',0.1),
plot(t,P(:,3),'p','linew',0.1),
plot(t,P(:,4),'.','linew',3),
plot(t,P(:,5),'--','linew',3),
plot(t,P(:,6),'-.','linew',3),
xlabel('\bf\it - - - - - t - - - - -'),
text(-0.1,1.04,'\bf P_k(t)'),
text(5,0.7, sprintf('%s%g%s%g', ...
'\bf\fontsize{12}\lambda\fontsize{10} = ',L,...
'\newline\fontsize{12}\mu\fontsize{10} = ', M)),
legend('P_0','P_1','P_2','P_3','P_4','P_5'),
set(gcf,'color', [0.7,0.7,0.7])

%% М-функция описания правых частей дифференциальных уравнений
function f = cmo11(t,P,A)
f = A*P;
```

Графическое представление решения системы (3.9) показано на рис. 3.4.

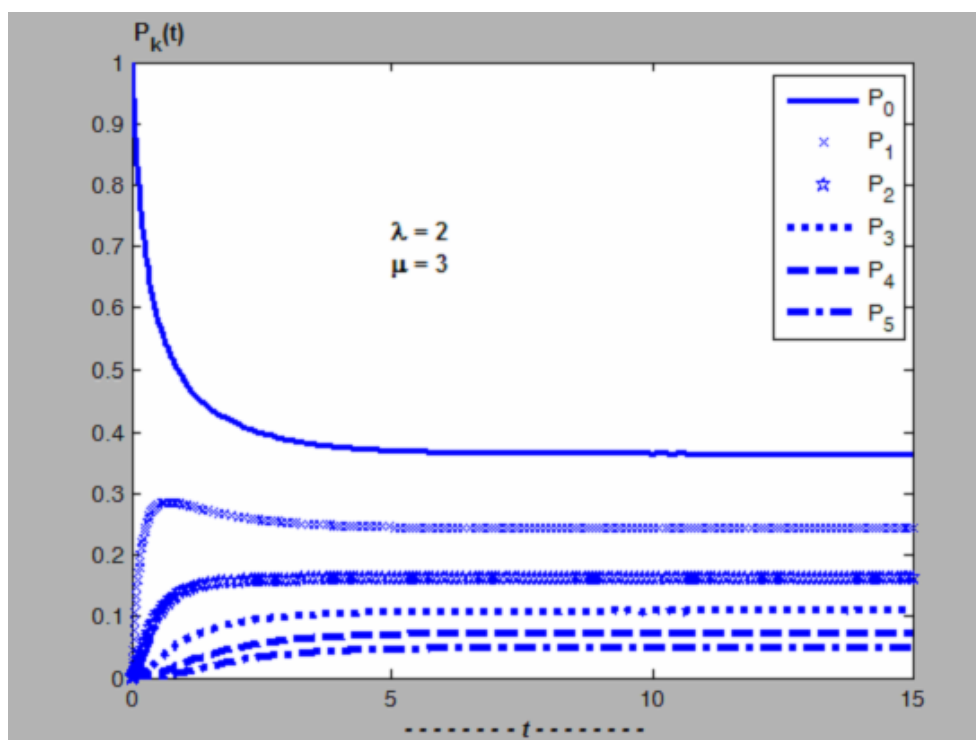


Рис. 3.4. Временные зависимости вероятностей системы M/M/1/5

Как видно из рис. 3.4, вероятности состояний с течением времени принимают практически установившиеся значения, которые достаточно близки к стационарным значениям вероятностей системы M/M/1 с бесконечным числом состояний (см. рис. 3.2). Конкретные значения вероятностей приведены в табл. 3.1.

Таблица 3.1

Вероятности состояний	Классическая M/M/1 с бесконечным числом состояний	Система M/M/1/5 с числом состояний 6
P_0	0,3333	0,3654
P_1	0,2222	0,2436
P_2	0,1481	0,1624
P_3	0,0988	0,1083
P_4	0,0658	0,0722
P_5	0,0439	0,0481
Сумма	0,921	1,0000

вероятностей		
--------------	--	--

В табл. 3.1 показана проверка нормировочного условия: сумма вероятностей для 6 состояний классической системы М/М/1 меньше единицы, так как нормировочное условие справедливо для полной группы событий, когда число состояний бесконечно. Для системы М/М/1 с числом состояний 6 нормировочное условие выполняется: сумма вероятностей равно единице.

3.2. Формула Литтла. Аналоги формулы Литтла

Прежде чем начать вывод формулы Литтла (из [4, 6]), отметим, что она справедлива для систем массового обслуживания (СМО) достаточно общего вида: СМО может быть одноканальной, многоканальной, системой с ограниченной или неограниченной очередью, системой с отказами, системой с произвольным входным потоком требований и произвольным законом их обслуживания. Формула Литтла устанавливает для стационарного режима СМО зависимость между интенсивностью λ входящего потока, временем T пребывания в системе одного требования и средним количеством требований \bar{N} , находящихся в системе [6, с.37]. Любая система массового обслуживания связана с двумя потоками событий: с потоком требований, прибывающих в систему, и с потоком требований, покидающих систему. При этом оба потока имеют одну и ту же интенсивность λ . В установившемся, предельном, стационарном режиме среднее число требований, прибывающих в СМО за единицу времени, равно среднему числу требований, покидающих её. При этом в качестве системы массового обслуживания может выступать система с очередью, система с отказами, система как накопитель, отображающий только очередь. В общем случае, когда рассматривается система с

очередью (с накопителем), то для нее будет получена собственно формула Литтла. В других случаях — системы с отказами и системы, представленные одним накопителем, — получаются аналоги формулы Литтла.

Примем следующие обозначения:

$\alpha(t)$ — число требований, прибывших в систему до момента t_1 , т. е. в промежутке $(0; t_1)$;

$\delta(t)$ — число требований, покидающих систему до момента времени t_1 .

Момент времени t_1 может быть любым и поэтому сами функции $\alpha(t)$ и $\delta(t)$ обозначены от аргумента t .

Введенные функции $\alpha(t)$ и $\delta(t)$ выражают собой вероятностные процессы во времени. Они являются случайными и меняются скачкообразно (увеличиваются на одну единицу) в момент прихода требований в систему и ухода требований из системы. Типовой характер изменения этих функций показан на рис. 3.5, где заштрихованная область соответствует области, ограниченной кривыми $\alpha(t)$ и $\delta(t)$. Вертикальные штриховые линии соответствуют произвольному моменту времени t_1 и моменту окончания T исследуемого процесса.

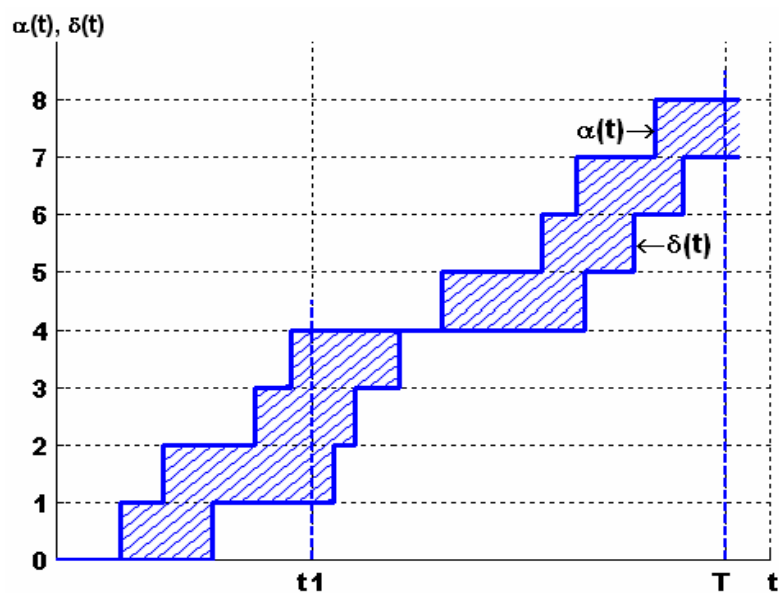


Рис. 3.5. Процесс прихода и ухода требований для произвольной СМО
(к выводу формулы Литтла)

Когда линии функций $\alpha(t)$ и $\delta(t)$ сливаются, то это означает, что в системе нет требований.

Для любого момента времени t (например, на рис. 3.5 такой момент времени обозначен как t_1) разность функций $\alpha(t)$ и $\delta(t)$ есть число требований, находящихся в системе, т. е.

$$N(t) = \alpha(t) - \delta(t). \quad (3.11)$$

Обозначим:

$\gamma(t)$ — площадь фигуры, ограниченная функциями $\alpha(t)$ и $\delta(t)$.

Очевидно, что с ростом t значение функции $\gamma(t)$ (площади) увеличивается. При этом отметим, что размерность $\gamma(t)$ есть произведение времени на число требований. На рис.3.5 площадь фигуры, ограниченная функциями $\alpha(t)$ и $\delta(t)$, заштрихована.

Обозначим:

$\lambda(t) = \frac{\alpha(t)}{t}$ — интенсивность поступления требований в систему до момента времени t .

Можно сказать, что интенсивность поступления требований — это число требований, поступивших в систему в единицу времени.

Определим время $T(t)$, проведенное одним требованием в системе, усредненное по всем требованиям, находящимся в системе до момента времени t .

Нам известно, что $\alpha(t)$ — число требований, поступивших в систему до момента времени t . Известно также, что $\gamma(t)$ — все требования, находившиеся в системе, до момента времени t . Выпишем размерности:

$\gamma(t)=[\text{требования} \cdot \text{с}]$ — требования на секунду, $\alpha(t) = [\text{требования}]$.

Тогда можно вычислить $T(t)$:

$$T(t) = \frac{\gamma(t)}{\alpha(t)}. \quad (3.12)$$

Найдем число требований $N(t)$, находящихся в системе до момента времени t :

$$N(t) = \frac{\gamma(t)}{t}. \quad (3.13)$$

Сопоставим соотношения по $N(t)$, $T(t)$, $\lambda(t)$:

$$N(t) = \frac{\gamma(t)}{t}, \quad T(t) = \frac{\gamma(t)}{\alpha(t)}, \quad \lambda(t) = \frac{\alpha(t)}{t}.$$

Из трех соотношений исключим время t и выполним следующие преобразования:

$$\lambda(t) = \frac{\alpha(t)}{t}, \Rightarrow t = \frac{\alpha(t)}{\lambda(t)},$$
$$N(t) = \frac{\gamma(t) \lambda(t)}{\alpha(t)} = \left\{ \gamma(t) = T(t) \alpha(t) \right\} = \frac{T(t) \alpha(t) \lambda(t)}{\alpha(t)} = T(t) \lambda(t).$$

Таким образом, количество требований в системе до момента времени t связаны с интенсивностью поступающих требований до момента времени t и временем пребывания одного требования до момента времени t соотношением

$$N(t) = \lambda(t) T(t). \quad (3.14)$$

Сделаем допущение о том, что в рассматриваемой системе массового обслуживания существуют конечные пределы:

$$\lim_{t \rightarrow \infty} \lambda(t) = \lambda, \quad (3.15)$$

$$\lim_{t \rightarrow \infty} T(t) = T. \quad (3.16)$$

При существовании указанных пределов получаем формулу Литтла:

$$N_{cp} = \lambda T, \quad (3.17)$$

где N_{cp} — среднее число требований в системе массового обслуживания, в которую требования поступают с постоянной интенсивностью λ и каждое из требований пребывает в этой системе T единиц времени (например, в секундах, как определяли выше).

Формула Литтла (3.17) справедлива для любой системы массового обслуживания, при любом характере потока требований, при любом распределении времени обслуживания, при любой дисциплине обслуживания [4, с.33]. Формула Литтла устанавливает, что среднее число N_{cp} требований в системе массового обслуживания равно произведению интенсивности поступления требований в систему на среднее время пребывания одного требования в системе.

Как видно из приведенного вывода формулы Литтла, этот вывод не зависит ни от вида вероятностного распределения входящего потока требований, ни от вида закона обслуживания требований, ни от числа обслуживающих приборов, ни от характера дисциплины обслуживания требований в системе. Более того, не зависит и от вида СМО. Например, в случае, когда ведется подсчет поступающих требований в очередь и исходящих из нее, то тогда получаем аналог формулы Литтла для очереди:

$$N_q = \lambda T_q, \quad (3.18)$$

где N_q — среднее число требований в очереди, T_q — среднее время ожидания в очереди, λ — интенсивность поступления требований в очередь или в систему обслуживания, для которой рассматривается очередь, индекс q — от английского queue (очередь).

Если система обслуживания не допускает очереди (система с отказами) и поэтому состоит только из обслуживающего прибора (приборов), то еще одним аналогом формулы Литтла будет являться равенство

$$N_s = \lambda T_s, \quad (3.19)$$

где N_s — среднее число требований в обслуживающем приборе (приборах), T_s — среднее время, проведенное требованием в обслуживающем приборе (приборах), λ — интенсивность поступления требований в обслуживающий прибор (приборы), индекс s — от английского serve (служить, обслуживать).

Очевидно, что для СМО, состоящей из очереди и обслуживающего прибора, выполняется равенство

$$T = T_q + T_s. \quad (3.20)$$

Среднее время, проведенное одним требованием в системе, складывается из времени, проведенное в очереди, и времени, проведенное в обслуживающем приборе.

3.3. Операционные характеристики системы М/М/1

3.3.1. Среднее число требований в системе М/М/1

Сначала подсчитаем среднее число \bar{N} требований, находящихся в системе. Количество требований в системе — это случайная дискретная величина, имеющая вероятностное распределение (3.8). Среднее число требований следует подсчитывать как математическое ожидание случайной дискретной величины:

$$\bar{N} = \sum_{k=0}^{\infty} k p_k = \sum_{k=1}^{\infty} k p_k = \sum_{k=1}^{\infty} k (1 - \rho) \rho^k, \quad (3.21)$$

где $\rho = \lambda / \mu$.

Вычисление суммы в (3.21) можно произвести непосредственным вычислением, раскрывая выражение суммы:

$$\begin{aligned}\bar{N} &= \sum_{k=1}^{\infty} k(1-\rho)\rho^k = \sum_{k=1}^{\infty} k\rho^k - \sum_{k=1}^{\infty} k\rho^{k+1} = \\ &= 1\rho + 2\rho^2 + 3\rho^3 + 4\rho^4 + \dots - 1\rho^2 - 2\rho^3 - 3\rho^4 - \dots = \\ &= \rho + \rho^2 + \rho^3 + \rho^4 + \dots = \frac{\rho}{1-\rho},\end{aligned}$$

где результирующая сумма ряда представляет собой убывающую геометрическую прогрессию со знаменателем ρ и первым членом также ρ . По условию

$$\rho < 1.$$

Таким образом, среднее число требований в системе М/М/1 определяется по формуле

$$\bar{N} = \frac{\rho}{1-\rho}. \quad (3.22)$$

Очевидно, что при $\rho \rightarrow 1$ число требований в системе неограниченно возрастает: $\bar{N} \rightarrow \infty$. При этом под количеством требований в системе понимается какое-то количество требований, находящихся в очереди, и одно требование, находящееся на обслуживании.

Рассмотрим определение \bar{N} с помощью приема, основанного на искусственном дифференцировании:

$$\bar{N} = \sum_{k=0}^{\infty} kp_k = \sum_{k=0}^{\infty} k(1-\rho)\rho^k = (1-\rho) \sum_{k=0}^{\infty} k\rho^k = (1-\rho)\rho \frac{d}{d\rho} \sum_{k=0}^{\infty} \rho^k. \quad (3.23)$$

В (3.23) сумма представляет собой убывающую геометрическую прогрессию, т.к. $\rho < 1$. Сумма членов убывающей геометрической прогрессии равна

$$\sum_{k=0}^{\infty} \rho^k = \frac{1}{1-\rho}, \quad (3.24)$$

где ρ — знаменатель геометрической прогрессии.

С учетом (3.24) и (3.23) получим

$$\bar{N} = (1 - \rho) \rho \frac{d}{d\rho} \sum_{k=0}^{\infty} \rho^k = (1 - \rho) \rho \frac{d}{d\rho} \left(\frac{1}{1 - \rho} \right) = (1 - \rho) \rho \frac{1}{(1 - \rho)^2} = \frac{\rho}{(1 - \rho)}.$$

Получен тот же результат, что и (3.22).

3.3.2. Дисперсия числа требований в системе М/М/1

Найдем дисперсию D_N среднего количества требований в системе М/М/1. По определению дисперсия дискретной случайной величины — это математическое ожидание квадрата отклонения случайной величины от своего математического ожидания [11]. В качестве случайной величины выступает число k требований в системе М/М/1. При этом если случайная величина имеет данное вероятностное распределение, то и функция от этой случайной величины будет иметь то же самое вероятностное распределение [11]. Поэтому можем записать следующее выражение для расчета дисперсии D_N :

$$D_N = \sum_{k=0}^{\infty} (k - \bar{N})^2 p_k,$$

где \bar{N} определяется по (3.22), p_k — по (3.8). Получим выражение для D_N :

$$\begin{aligned} D_N &= \sum_{k=0}^{\infty} (k - \bar{N})^2 p_k = \sum_{k=0}^{\infty} (k - \bar{N})^2 (1 - \rho) \rho^k = (1 - \rho) \sum_{k=0}^{\infty} (k^2 - 2k\bar{N} + \bar{N}^2) \rho^k = \\ &= (1 - \rho) \left(\sum_{k=0}^{\infty} k^2 \rho^k - 2\bar{N} \sum_{k=0}^{\infty} k \rho^k + \bar{N}^2 \sum_{k=0}^{\infty} \rho^k \right). \end{aligned} \quad (3.25)$$

Вычислим каждую из трех сумм в последнем выражении. Третья сумма представляет собой убывающую геометрическую прогрессию, поэтому ее вычисляем по формуле (3.24), а \bar{N} по (3.22):

$$\bar{N}^2 \sum_{k=0}^{\infty} \rho^k = \bar{N}^2 \frac{1}{1 - \rho} = \frac{\rho^2}{(1 - \rho)^2} \cdot \frac{1}{1 - \rho} = \frac{\rho^2}{(1 - \rho)^3}. \quad (3.26)$$

Для вычисления второй суммы используем прием искусственного дифференцирования:

$$\begin{aligned} 2\bar{N} \sum_{k=0}^{\infty} k \rho^k &= 2\bar{N} \rho \frac{d}{d\rho} \sum_{k=0}^{\infty} \rho^k = 2\bar{N} \rho \frac{d}{d\rho} \left(\frac{1}{1-\rho} \right) = 2\bar{N} \rho \frac{1}{(1-\rho)^2} = \\ &= 2 \frac{\rho}{1-\rho} \cdot \frac{\rho}{(1-\rho)^2} = \frac{2\rho^2}{(1-\rho)^3}. \end{aligned} \quad (3.27)$$

Для вычисления первой суммы в (3.25) дважды применим прием искусственного дифференцирования:

$$\begin{aligned} \sum_{k=0}^{\infty} k^2 \rho^k &= \rho \frac{d}{d\rho} \sum_{k=0}^{\infty} k \rho^k = \rho \frac{d}{d\rho} \left(\rho \frac{d}{d\rho} \sum_{k=0}^{\infty} \rho^k \right) = \\ &= \rho \frac{d}{d\rho} \left(\rho \frac{d}{d\rho} \left(\frac{1}{1-\rho} \right) \right) = \rho \frac{d}{d\rho} \left(\frac{\rho}{(1-\rho)^2} \right) = \rho \frac{(1-\rho)^2 + 2\rho(1-\rho)}{(1-\rho)^4} = \\ &= \frac{\rho(1-\rho) + 2\rho^2}{(1-\rho)^3}. \end{aligned} \quad (3.28)$$

Выражения (3.26)-(3.28) используем для расчета D_N по (3.25):

$$\begin{aligned} D_N &= (1-\rho) \left(\frac{\rho(1-\rho) + 2\rho^2}{(1-\rho)^3} - \frac{2\rho^2}{(1-\rho)^3} + \frac{\rho^2}{(1-\rho)^3} \right) = \\ &= \frac{\rho - \rho^2 + 2\rho^2 - \rho^2}{(1-\rho)^2} = \frac{\rho}{(1-\rho)^2}. \end{aligned}$$

Таким образом, дисперсия числа требований в системе М/М/1 будет вычисляться по следующей формуле:

$$D_N = \frac{\rho}{(1-\rho)^2}. \quad (3.29)$$

Также как и для среднего числа требований в системе дисперсия $D_N \rightarrow \infty$ при стремлении $\rho \rightarrow 1$.

По известной величине дисперсии может быть найдено среднее квадратическое отклонение (стандартное отклонение) σ_N числа требований в системе:

$$\sigma_N = +\sqrt{D_N} = \frac{+\sqrt{\rho}}{1-\rho}. \quad (3.30)$$

Для дисперсии часто используют обозначение, выраженное через среднее квадратическое отклонение:

$$D_N = \sigma_N^2. \quad (3.31)$$

3.3.3. Вероятность нахождения в системе М/М/1 не менее k требований

Обозначим вероятность того, что в системе имеется, по крайней мере, k требований, как $P(\geq k)$. Искомая характеристика определяется в соответствии с теоремой сложения вероятностей несовместных событий:

$$\begin{aligned} P(\geq k) &= \sum_{i=k}^{\infty} p_i = \sum_{i=k}^{\infty} (1-\rho)\rho^i = \sum_{i=k}^{\infty} (\rho^i - \rho^{i+1}) = \\ &= \rho^k - \rho^{k+1} + \rho^{k+1} - \rho^{k+2} + \rho^{k+2} - \rho^{k+3} + \rho^{k+3} - \dots = \rho^k. \end{aligned} \quad (3.32)$$

Получили, что вероятность того, что число требований в системе М/М/1 превысит некоторое граничное значение k , описывается убывающей геометрической прогрессией, которая зависит от этого граничного числа [4]. Вывод может быть произведен с использованием формулы убывающей геометрической прогрессии с первым членом ρ^k и знаменателем ρ :

$$P(\geq k) = \sum_{i=k}^{\infty} p_i = \sum_{i=k}^{\infty} (1-\rho)\rho^i = (1-\rho) \sum_{i=k}^{\infty} \rho^i = (1-\rho) \frac{\rho^k}{(1-\rho)} = \rho^k.$$

Получен тот же результат.

Можно подсчитать вероятность $P(< k)$ нахождения в системе не более k требований, как дополнение к вероятности $P(\geq k)$:

$$P(< k) = 1 - P(\geq k) = 1 - \rho^k.$$

На рис. 3.6 показан пример изменения операционной характеристики $P(\geq k)$ с ростом k при $\rho = 0,5$.

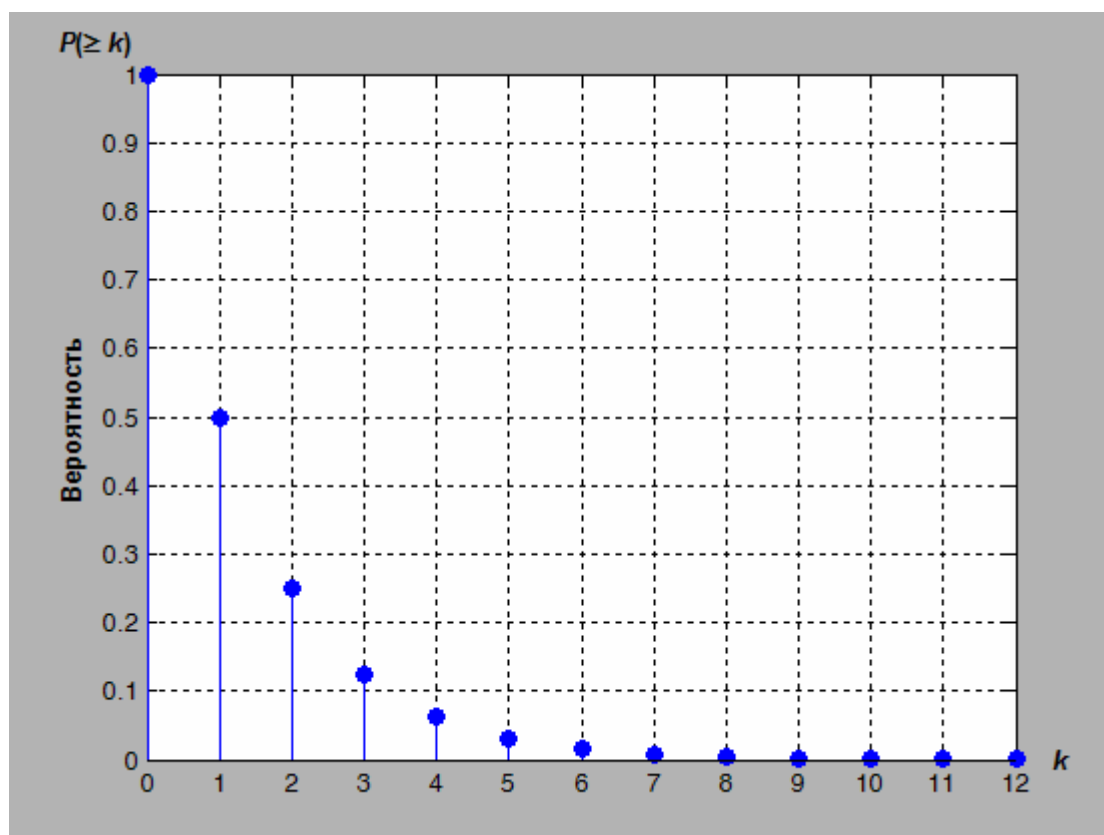


Рис. 3.6. Изменение операционной характеристики $P(\geq k)$ с ростом k

Ряд значений вероятности $P(\geq k)$ при $\rho = 0,5$ образует геометрическую прогрессию со знаменателем 0,5. Некоторые значения этого ряда приведены в табл. 3.2.

Таблица 3.2

k	0	1	2	3	4	5	6	7
$P(\geq k)$	1	0,5	0,25	0,125	0,0625	0,0313	0,0156	0,0078

3.3.4. Среднее число требований, находящихся в очереди системы М/М/1

Требование попадает в очередь, если прибор обслуживания занят. Следовательно, в очереди будет находиться такое количество требований, которое находится в системе за минусом одного обслуживаемого

требования. Очередь начинает формироваться, когда в системе находится, по крайней мере, два требования, одно из которых находится на обслуживании. Как всякое среднее оно может быть вычислено через математическое ожидание и в данном случае через математическое ожидание случайной дискретной величины k — числа требований. Обозначим искомую операционную характеристику как \bar{N}_q . Вычислим \bar{N}_q :

$$\begin{aligned}\bar{N}_q &= \sum_{k=2}^{\infty} (k-1)p_k = \sum_{k=2}^{\infty} (k-1)(1-\rho)\rho^k = (1-\rho) \sum_{k=2}^{\infty} (k-1)\rho^k = \\ &= (1-\rho) \sum_{k=2}^{\infty} k\rho^k - (1-\rho) \sum_{k=2}^{\infty} \rho^k = (1-\rho)\rho \frac{d}{d\rho} \sum_{k=2}^{\infty} \rho^k - (1-\rho) \frac{\rho^2}{1-\rho} = \\ &= (1-\rho)\rho \frac{d}{d\rho} \left(\frac{\rho^2}{1-\rho} \right) - \rho^2 = (1-\rho)\rho \frac{2\rho(1-\rho) + \rho^2}{(1-\rho)^2} - \rho^2 = \\ &= \frac{2\rho^2 - 2\rho^3 + \rho^3}{1-\rho} - \rho^2 = \frac{2\rho^2 - \rho^3 - \rho^2(1-\rho)}{1-\rho} = \frac{2\rho^2 - \rho^3 - \rho^2 + \rho^3}{1-\rho} = \frac{\rho^2}{1-\rho}.\end{aligned}$$

Таким образом, получили следующее выражение для \bar{N}_q :

$$\bar{N}_q = \frac{\rho^2}{1-\rho}, \quad (3.33)$$

где $\rho = \frac{\lambda}{\mu} < 1$.

Среднее число требований, находящихся в очереди, еще называют средней длиной очереди, в данном случае средней длиной очереди в системе М/М/1. Как видно из выражения (3.33) средняя длина очереди (среднее число требований в очереди) будет неограниченно увеличиваться при $\rho \rightarrow 1$.

3.3.5. Среднее время пребывания требования в системе М/М/1

Среднее время T пребывания одного требования в системе М/М/1 определим по формуле Литтла (3.17) с учетом (3.22):

$$\bar{N} = \lambda T, \Rightarrow T = \frac{\bar{N}}{\lambda},$$
$$T = \frac{\bar{N}}{\lambda} = \frac{\rho}{(1-\rho)\lambda} = \frac{\lambda}{\mu(1-\rho)\lambda} = \frac{1}{\mu(1-\rho)}. \quad (3.34)$$

Из (3.34) видно, что при $\rho \rightarrow 1$ среднее время пребывания в системе неограниченно увеличивается.

3.3.6. Среднее время ожидания в очереди системы М/М/1

Среднее время ожидания в очереди одного требования определим по аналогу формулы Литтла (3.18) с учетом (3.33):

$$\bar{N}_q = \lambda T_q, \Rightarrow T_q = \frac{\bar{N}_q}{\lambda} = \frac{\rho^2}{(1-\rho)\lambda}. \quad (3.35)$$

Очевидно, что среднее время ожидания требования в очереди системы М/М/1 будет неограниченно увеличиваться при $\rho \rightarrow 1$.

3.3.7. Вероятность того, что вновь поступающее требование попадет в очередь системы М/М/1

Требование попадает в очередь системы М/М/1, если обслуживающий прибор занят. Следовательно, в системе находится как минимум одно требование (которое на обслуживание). Это означает, что надо определить вероятность P_q нахождения в системе не менее 2-х требований. Для этого используем выражение (3.32):

$$P_q = P(\geq 2) = \sum_{i=2}^{\infty} p_i = \rho^2.$$

Таким образом, имеем

$$P_q = \rho^2 = \left(\frac{\lambda}{\mu} \right)^2. \quad (3.36)$$

Выражение (3.36) можно получить на основе формулы убывающей геометрической прогрессии:

$$\begin{aligned} P_q = P(\geq 2) &= \sum_{i=2}^{\infty} p_i = \sum_{i=2}^{\infty} (1-\rho)\rho^i = \sum_{i=2}^{\infty} (\rho^i - \rho^{i+1}) = \sum_{i=2}^{\infty} \rho^i - \sum_{i=2}^{\infty} \rho^{i+1} = \\ &= \frac{\rho^2}{1-\rho} - \frac{\rho^3}{1-\rho} = \frac{\rho^2(1-\rho)}{1-\rho} = \rho^2. \end{aligned}$$

3.3.8. Функция и плотность распределения времени пребывания требования в системе М/М/1

Получим функцию распределения $W(\tau)$ и плотность распределения $w(\tau)$ времени пребывания одного требования в системе М/М/1. Будем предполагать, что очередь в системе формируется по дисциплине FIFO (первым пришел, первым вышел). Пусть τ означает время, в течение которого требование вынуждено оставаться в системе до завершения обслуживания. При наличии очереди вновь поступившее требование по счету будет каким-то $n+1$. Тогда все время пребывания в системе для $(n+1)$ -го требования будет состоять из следующих отрезков времени:

$$\tau = t_{1s} + t_2 + \dots + t_n + t_{n+1}, \quad (3.37)$$

где t_{1s} — время, необходимое для завершения обслуживания требования, которое уже находится в приборе обслуживания, t_2, \dots, t_n — отрезки времени, которые потребуются для обслуживания 2-го, ..., n -го требований, которые в момент прибытия $(n+1)$ -го требования находятся в

очереди (в накопителе или блоке ожидания), t_{n+1} — длина отрезка времени, по истечении которого будет обслужено $(n+1)$ -е требование [3, с.214]. В системе М/М/1 поступающие требования независимы друг от друга и обслуживаются по экспоненциальному закону и, следовательно, τ есть сумма $(n+1)$ независимых величин, распределенных по экспоненциальному закону. Введем в рассмотрение $P(\tau | n+1)$ условную вероятность распределения величины τ , когда перед только что поступившим в систему требованием окажется ровно n ранее поступивших требований. В общем случае n может принимать значение от 0 до бесконечности. Поступающие в систему М/М/1 требования имеют распределение вероятностей p_k , определяемое выражением (3.8). Условная вероятность распределения величины τ есть распределение Эрланга, определяемое выражением (1.102) с заменой λ на μ . Вероятность поступления в систему точно n требований есть независимое событие. Тогда функция распределения величины τ может быть найдена по формуле полной вероятности, учитывая связь функции распределения со своей плотностью:

$$\begin{aligned} W(\tau) &= \sum_{n=0}^{\infty} P(\tau | n+1) p_n = \sum_{n=0}^{\infty} \int_0^{\tau} \frac{\mu(\mu t)^n}{n!} e^{-\mu t} dt (1-\rho) \rho^n = \\ &= \int_0^{\tau} \left[\sum_{n=0}^{\infty} \frac{\mu(\mu t)^n}{n!} e^{-\mu t} (1-\rho) \rho^n \right] dt. \end{aligned}$$

Целесообразно сначала получить плотность распределения $w(\tau)$, для чего продифференцируем по верхнему пределу интегрирования выражения в записи функции распределения $W(\tau)$:

$$\begin{aligned} w(\tau) &= \frac{dW(\tau)}{d\tau} = \sum_{n=0}^{\infty} \frac{\mu(\mu\tau)^n e^{-\mu\tau} (1-\rho)\rho^n}{n!} = \mu e^{-\mu\tau} (1-\rho) \sum_{n=0}^{\infty} \frac{(\mu\tau\rho)^n}{n!} = \\ &= \mu e^{-\mu\tau} (1-\rho) \sum_{n=0}^{\infty} \frac{(\mu\tau\lambda / \mu)^n}{n!} = \mu e^{-\mu\tau} (1-\rho) \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!}. \end{aligned}$$

В полученном выражении последняя сумма есть разложение экспоненты с показателем $\lambda\tau$. Поэтому имеем

$$\begin{aligned} w(\tau) - (1-\rho)\mu e^{-\mu\tau} e^{\lambda\tau} &= \mu(1-\rho)e^{-(\mu-\lambda)\tau} = \\ &= \mu(1-\rho)e^{-(\mu-\lambda/\mu)\tau} = \mu(1-\rho)e^{-\mu(1-\rho)\tau}. \end{aligned}$$

Таким образом, получили следующую плотность распределения случайной величины τ :

$$w(\tau) = \mu(1-\rho)e^{-\mu(1-\rho)\tau}, \quad \tau \geq 0. \quad (3.38)$$

Выражение (3.38) соответствует экспоненциальному распределению с параметром $\mu(1-\rho)$. Соответствующая функция распределения $W(\tau)$ находится путем интегрирования функции плотности (3.38):

$$W(\tau) = \int_0^{\tau} \mu(1-\rho)e^{-\mu(1-\rho)t} dt = -e^{-\mu(1-\rho)t} \Big|_0^{\tau} = -e^{-\mu(1-\rho)\tau} + 1$$

или

$$W(\tau) = 1 - e^{-\mu(1-\rho)\tau}, \quad \tau \geq 0. \quad (3.39)$$

Функция распределения (3.39) означает вероятность того, что время пребывания требования в системе не превысит величины τ . Среднее значение $M(\tau)$ экспоненциального распределения (3.38), (3.39) будет иметь вид

$$M(\tau) = \frac{1}{\mu(1-\rho)}. \quad (3.40)$$

3.4. Система М/М/1 с отказами

3.4.1. Вероятности состояний системы М/М/1 с отказами

Проанализируем систему обслуживания М/М/1 для случая, когда требование, заставшее систему занятой, сразу же покидает ее. Для любого момента времени $t \geq 0$ система оказывается либо в состоянии S_0 (канал свободен), либо в состоянии S_1 (канал занят). Переход из S_1 в S_0 осуществляется, как только очередное обслуживание завершится [16]. Переход из S_0 в S_1 связан с появлением требования и немедленным началом его обслуживания. Размеченный граф состояний для системы М/М/1 с отказами показан на рис. 3.7.

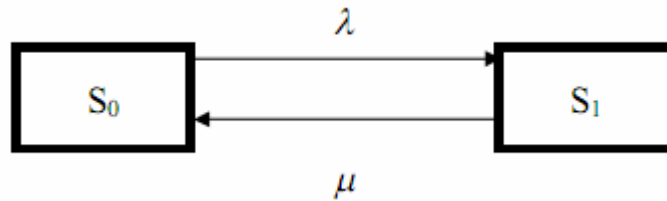


Рис. 3.7. Граф состояний для системы М/М/1 с отказами

Изменение состояний во времени для системы М/М/1 с отказами может быть найдено из решения дифференциальных уравнений процесса размножения и гибели (2.10) с начальными условиями типа (2.12) и с коэффициентами $\lambda_0 = \lambda$, $\mu_1 = \mu$. Запишем уравнение относительно $P_0(t)$:

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t). \quad (3.41)$$

Из нормировочного условия выразим $P_1(t)$ через $P_0(t)$:

$$\sum_{k=0}^1 P_k(t) = 1, \Rightarrow P_0(t) + P_1(t) = 1, \quad P_1(t) = 1 - P_0(t). \quad (3.42)$$

С учетом (3.42) запишем уравнение (3.41) в виде

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda P_0(t) + \mu[1 - P_0(t)], \\ \frac{dP_0(t)}{dt} &= -(\lambda + \mu)P_0(t) + \mu. \end{aligned} \quad (3.43)$$

Полученное уравнение (3.43) — это неоднородное дифференциальное уравнение. Его интегрирование следует проводить при естественных начальных условиях [16, с.190], поскольку в начальный момент система свободна:

$$P_0(0) = 1, \quad P_1(0) = 0. \quad (3.44)$$

Сначала решим соответствующее однородное уравнение:

$$\frac{dP_0(t)}{dt} = -(\lambda + \mu)P_0(t), \quad \frac{dP_0(t)}{P_0(t)} = -(\lambda + \mu)dt, \Rightarrow \ln P_0(t) = -(\lambda + \mu)t + \ln C,$$

где C — постоянная интегрирования. После потенцирования логарифмического уравнения получим

$$P_0(t) = Ce^{-(\lambda + \mu)t}. \quad (3.45)$$

Затем применим метод вариации произвольных постоянных:

$$\begin{aligned} C = C(t), \Rightarrow P_0(t) &= C(t)e^{-(\lambda + \mu)t}, \quad \frac{dP_0(t)}{dt} = \frac{dC(t)}{dt}e^{-(\lambda + \mu)t} - (\lambda + \mu)C(t)e^{-(\lambda + \mu)t}, \\ \frac{dC(t)}{dt}e^{-(\lambda + \mu)t} - (\lambda + \mu)C(t)e^{-(\lambda + \mu)t} &= -(\lambda + \mu)C(t)e^{-(\lambda + \mu)t} + \mu, \quad \frac{dC(t)}{dt} = \mu e^{(\lambda + \mu)t}, \\ C(t) &= \frac{\mu}{\lambda + \mu}e^{(\lambda + \mu)t} + c, \quad P_0(t) = \left[\frac{\mu}{\lambda + \mu}e^{(\lambda + \mu)t} + c \right] e^{-(\lambda + \mu)t}, \\ P_0(t) &= \frac{\mu}{\lambda + \mu} + ce^{-(\lambda + \mu)t}. \end{aligned}$$

Используя начальные условия (3.44), определим постоянную интегрирования c :

$$t = 0: \quad 1 = \frac{\mu}{\lambda + \mu} + c, \quad c = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda + \mu - \mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}.$$

Таким образом, получаем решение для вероятности состояния $P_0(t)$:

$$P_0(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}e^{-(\lambda + \mu)t}. \quad (3.46)$$

Из нормировочного условия (3.42) определяем $P_1(t)$:

$$P_1(t) = 1 - P_0(t) = 1 - \frac{\mu}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} =$$

$$= \frac{\lambda + \mu - \mu - \lambda e^{-(\lambda + \mu)t}}{\lambda + \mu} = \frac{\lambda + \lambda e^{-(\lambda + \mu)t}}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu} [1 - e^{-(\lambda + \mu)t}].$$

Решение для $P_1(t)$ имеет вид

$$P_1(t) = \frac{\lambda}{\lambda + \mu} [1 - e^{-(\lambda + \mu)t}]. \quad (3.47)$$

Стационарный режим работы системы М/М/1 с отказами (стационарные вероятности p_0, p_1) легко может быть получен из (3.46), (3.47) как предельное выражение при $t \rightarrow \infty$:

$$p_0 = \frac{\mu}{\lambda + \mu}, \quad (3.48)$$

$$p_1 = \frac{\lambda}{\lambda + \mu}. \quad (3.49)$$

3.4.2. Операционные характеристики системы М/М/1 с отказами

Для системы М/М/1 с отказами интерес представляют следующие операционные характеристики (для стационарного режима): *вероятность отказа в обслуживании* $P_{отк}$, *относительная пропускная способность* Q , *абсолютная пропускная способность* A .

Вероятность $P_{отк}$ характеризует долю требований, получающих отказ. Требование, поступающее в систему, получает отказ в том случае, когда прибор (канал) обслуживания занят, т. е. когда в системе уже находится одно требование, а сама система пребывает в состоянии S_1 . Вероятность пребывания в системе одного требования определяется по (3.49). Поэтому вероятность отказа равна

$$P_{отк} = p_1 = \frac{\lambda}{\lambda + \mu} = \frac{\lambda}{\mu(1 + \lambda/\mu)} = \frac{\rho}{1 + \rho}. \quad (3.50)$$

Относительная пропускная способность системы с отказами — это вероятность того, что требование (заявка), поступившее в систему, будет обслужено. Для системы М/М/1 с отказами полная группа событий образуется из вероятности обслуживания требований и вероятности отказа. Поэтому

$$Q = 1 - P_{отк} = 1 - \frac{\lambda}{\lambda + \mu} = \frac{\lambda + \mu - \lambda}{\lambda + \mu} = \frac{\mu}{\lambda + \mu} = \frac{\mu}{\mu(1 + \lambda/\mu)} = \frac{1}{1 + \rho}. \quad (3.51)$$

Как видно из (3.50), (3.51) для потоков, у которых большие значения интенсивностей ($\lambda \gg \mu$), вероятность отказа системы стремится к единице, а относительная пропускная способность падает до нуля.

Абсолютная пропускная способность системы с отказами — это среднее число требований, обслуженных в единицу времени. Число требований, поступающих в систему в единицу времени, характеризуется величиной интенсивности λ , а вероятность обслуживания есть относительная пропускная способность. Поэтому

$$A = \lambda Q = \frac{\lambda}{1 + \rho}. \quad (3.52)$$

Величины $P_{отк}$, Q являются безразмерными, а величина A измеряется в c^{-1} .

3.5. Система М/М/1/К: конечный накопитель

3.5.1. Вероятности состояний системы М/М/1/К

Произведем анализ системы М/М/1/К, т. е. системы, для которой ограничено число ожидающих требований или ограничен объем накопителя (ограничена длина очереди). В обозначении системы буква **К** указывает на то, что в системе могут находиться самое большее **К** требований, включая требование, находящееся в приборе обслуживания. Если поступает **К+1** требование, то оно покидает систему. Размеченный граф состояний системы показан на рис. 3.8.

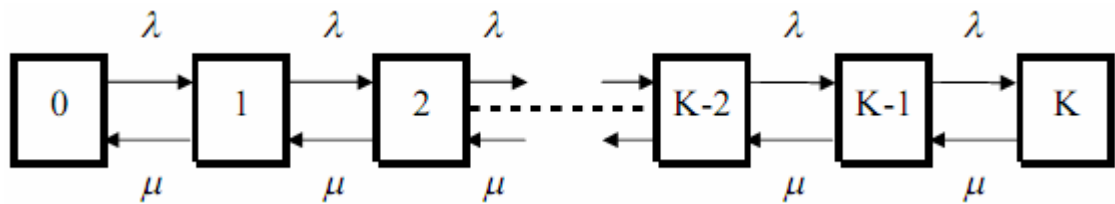


Рис. 3.8. Граф состояний системы М/М/1/К с конечным накопителем

В произвольный момент времени система М/М/1/К оказывается в одном из состояний 0 (канал обслуживания свободен, очереди нет), 1 (канал обслуживания занят, очереди нет), 2 (канал обслуживания занят, в очереди одно требование), ..., **К** (канал обслуживания занят, в очереди **К**–1 требований). Для такой системы устанавливаются следующие параметры процесса размножения и гибели:

$$\lambda_k = \begin{cases} \lambda = const, & k < K; \\ 0, & k \geq K; \end{cases} \quad (3.53)$$

$$\mu_k = \mu = const, \quad k = \overline{1, K}.$$

Как и в случае системы М/М/1 с неограниченной очередью стационарные вероятности состояний p_k могут быть определены по формулам (2.46), (2.47) с учетом (3.53):

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu} = p_0 \underbrace{\frac{\lambda}{\mu} \cdot \frac{\lambda}{\mu} \cdot \dots \cdot \frac{\lambda}{\mu}}_k = p_0 \left(\frac{\lambda}{\mu} \right)^k = p_0 \rho^k, \quad k = 0, 1, 2, \dots, K;$$

$$p_0 = \frac{1}{1 + \sum_{k=1}^K \prod_{i=0}^{k-1} \frac{\lambda}{\mu}} = \frac{1}{1 + \sum_{k=1}^K \left(\frac{\lambda}{\mu}\right)^k} = \frac{1}{1 + \sum_{k=1}^K \rho^k}.$$

В выражении для p_0 сумма представляет собой сумму конечной геометрической прогрессии. Напомним, чему она равна:

$$\sum_{j=0}^n a_j = \sum_{j=0}^n a_0 r^j = a_0 \frac{1 - r^{n+1}}{1 - r}, \quad (3.54)$$

где a_0 — первое слагаемое геометрической прогрессии, r — знаменатель геометрической прогрессии [17, с.30].

Для рассматриваемого случая можно записать p_0 в виде

$$p_0 = \left[1 + \sum_{k=1}^K \rho^k \right]^{-1} = \left[\sum_{k=0}^K \rho^k \right]^{-1}. \quad (3.55)$$

Сопоставляя (3.55) и (3.54), можем отметить, что для (3.55) знаменатель геометрической прогрессии $r=\rho$, $a_0=1$. Тогда получаем

$$p_0 = \left[\sum_{k=0}^K \rho^k \right]^{-1} = \left[\frac{1 - \rho^{K+1}}{1 - \rho} \right]^{-1} = \frac{1 - \rho}{1 - \rho^{K+1}}. \quad (3.56)$$

Таким образом, для системы М/М/1/К в стационарном режиме имеют место следующие вероятности состояний:

$$p_k = p_0 \rho^k, \quad k = 0, 1, 2, \dots, K; \quad (3.57)$$

$$p_0 = \frac{1 - \rho}{1 - \rho^{K+1}}. \quad (3.58)$$

Кроме того, имеет место равенство

$$p_k = 0, \quad k > K. \quad (3.59)$$

Приведем вероятности состояний для системы с удалением заблокированных вызовов ($K = 1$):

$$p_0 = \frac{1 - \rho}{1 - \rho^2} = \frac{1 - \rho}{(1 - \rho)(1 + \rho)} = \frac{1}{1 + \rho};$$

$$p_1 = p_0 \rho^1 = \frac{\rho}{1 + \rho}.$$

Полученные соотношения представим в виде, как это сделано в [4, с.122]:

$$p_k = \begin{cases} \frac{1}{1 + \rho} = \frac{1}{1 + \lambda / \mu}, & k = 0; \\ \frac{\rho}{1 + \rho} = \frac{\lambda / \mu}{1 + \lambda / \mu}, & k = 1 = K; \\ 0, & k > 1. \end{cases} \quad (3.60)$$

В системе с удалением заблокированных вызовов может находиться только одно требование. Требования, которые попытаются поступить в систему, когда она находится в состоянии обслуживания, будут из системы удалены. Вероятность поступления в систему больше одного требования равна нулю.

3.5.2. Операционные характеристики системы М/М/1/К

Вычислим для системы М/М/1/К среднюю длину N_q очереди. Она определяется как математическое ожидание дискретной случайной величины с возможными значениями 1, 2, ..., $K-1$, вероятности которых соответственно равны p_2, p_3, \dots, p_K , где K — допустимое число требований в системе (в очереди и в обслуживающем приборе). Тогда

$$N_q = \sum_{k=2}^K (k-1)p_k = \sum_{k=1}^K kp_{k+1} = p_0 \rho \sum_{k=1}^K k \rho^k. \quad (3.61)$$

Сумма в (3.61) может быть найдена по формуле конечной геометрической прогрессии (3.54). Раскроем эту сумму в (3.61):

$$\begin{aligned}
 \sum_{k=1}^K k \rho^k &= \rho \frac{d}{d\rho} \left(\sum_{k=1}^K \rho^k \right) = \rho \frac{d}{d\rho} \left(\frac{\rho(1-\rho^K)}{1-\rho} \right) = \\
 &= \rho \frac{d}{d\rho} \left(\frac{\rho(1-\rho^K)}{1-\rho} \right) = \rho \cdot \frac{(1-\rho^K - K\rho^K)(1-\rho) + \rho(1-\rho^K)}{(1-\rho)^2} = \\
 &= \frac{\rho(1-\rho^K - K\rho^K - \rho + \rho^{K+1} + K\rho^{K+1} + \rho - \rho^{K+1})}{(1-\rho)^2} = \\
 &= \frac{\rho[1 - (K+1)\rho^K + K\rho^{K+1}]}{(1-\rho)^2}.
 \end{aligned}$$

При раскрытии суммы в (3.61) использовали формулу конечной геометрической прогрессии с первым членом ρ и знаменателем ρ . Подставляя значение ρ_0 из (3.58), получим следующее выражение для N_q :

$$N_q = \frac{1-\rho}{1-\rho^{K+1}} \cdot \frac{\rho^2[1 - (K+1)\rho^K + K\rho^{K+1}]}{(1-\rho)^2} = \frac{\rho^2[1 - (K+1)\rho^K + K\rho^{K+1}]}{(1-\rho^{K+1})(1-\rho)}.$$

Таким образом, средняя длина очереди для системы М/М /1/К равна

$$N_q = \frac{\rho^2[1 - (K+1)\rho^K + K\rho^{K+1}]}{(1-\rho^{K+1})(1-\rho)}. \quad (3.62)$$

Аналогичным образом, как это сделано для средней длины очереди, можно определить среднее число N требований в системе М/М/1/К. Для этого следует выполнить преобразования выражения

$$N = \sum_{k=1}^K k p_k = \sum_{k=1}^K k p_0 \rho^k.$$

Определим вероятность P_q наличия очереди в системе М/М/1/К. Величина P_q означает вероятность того, что в системе будет не менее 2-х требований и не более K требований. Поэтому величину P_q вычисляем как сумму вероятностей:

$$\begin{aligned} P_q &= \sum_{k=2}^K p_k = p_0 \sum_{k=2}^K \rho^k = p_0 \cdot \frac{\rho^2(1 - \rho^{K-1})}{1 - \rho} = \frac{1 - \rho}{(1 - \rho^{K+1})} \cdot \frac{\rho^2(1 - \rho^{K-1})}{1 - \rho} = \\ &= \frac{\rho^2(1 - \rho^{K-1})}{(1 - \rho^{K+1})}. \end{aligned}$$

Таким образом, вероятность наличия очереди или вероятность попадания в очередь равна

$$P_q = \frac{\rho^2(1 - \rho^{K-1})}{(1 - \rho^{K+1})}. \quad (3.63)$$

Вероятность $P_{отк}$ отказа в обслуживании определяется в том случае, когда в системе одно требование находится в обслуживающем приборе и $K-1$ требований в накопителе (в очереди):

$$P_{отк} = p_K = p_0 \rho^K, \quad (3.64)$$

где

$$p_0 = \frac{1 - \rho}{(1 - \rho^{K+1})}.$$

Зная вероятность отказа, можно определить относительную пропускную способность q системы М/М/1/К:

$$\begin{aligned} q = 1 - P_{отк} &= 1 - p_0 \rho^K = 1 - \frac{(1 - \rho)\rho^K}{(1 - \rho^{K+1})} = \frac{1 - \rho^{K+1} - \rho^K + \rho^{K+1}}{(1 - \rho^{K+1})} = \frac{1 - \rho^K}{1 - \rho^{K+1}}. \end{aligned} \quad (3.65)$$

По известной относительной пропускной способности вычисляется абсолютная пропускная способность Q системы М/М/1/К:

$$Q = \lambda q = \lambda(1 - P_{отк}) = \frac{\lambda(1 - \rho^K)}{(1 - \rho^{K+1})}. \quad (3.66)$$

Определим такую операционную характеристику системы М/М/1/К как среднее время T_q пребывания в очереди. Величина T_q может быть получена, исходя из следующих рассуждений [16, с.197]. Если требование приходит в некоторый момент времени $t > 0$, то оно либо застанет систему свободной и сразу будет обслужено (вероятность этого события есть p_0), либо застанет систему занятой, но не имеющей очереди, и будет вынуждено ждать начала обслуживания в течение какого-то времени (его среднее значение равно $1/\mu$, а вероятность рассматриваемого события есть p_1), либо застанет очередь из одного требования, пришедшего ранее, и будет ожидать большее время (в среднем $2/\mu$ с вероятностью p_2) и т. д. Наконец, застав очередь длиной $K-1$ (плюс одно требование в обслуживающем приборе), требование останется не обслуженным и покинет систему. Это означает, что в качестве случайной величины выступает среднее время ожидания k/μ , где $k = 1, 2, \dots, K$, с вероятностями p_1, p_2, \dots, p_K . Величина T_q может быть определена как математическое ожидание случайной дискретной величины с известным распределением вероятностей p_1, p_2, \dots, p_K в соответствии с (3.57):

$$\begin{aligned}
 T_q &= \sum_{k=1}^K \frac{k p_k}{\mu} = \sum_{k=1}^K \frac{k p_0 \rho^k}{\mu} = \frac{p_0}{\mu} \rho \frac{d}{d\rho} \sum_{k=1}^K \rho^k = \frac{p_0}{\mu} \rho \frac{d}{d\rho} \left(\frac{\rho(1-\rho^K)}{1-\rho} \right) = \\
 &= \frac{p_0 \rho}{\mu} \cdot \frac{[1-\rho^K - K\rho^K](1-\rho) + \rho(1-\rho^K)}{(1-\rho)^2} = \\
 &= \frac{p_0 \rho}{\mu} \cdot \frac{1-\rho^K - K\rho^K - \rho + \rho^{K+1} + K\rho^{K+1} + \rho - \rho^{K+1}}{(1-\rho)^2} = \\
 &= \frac{p_0 \rho}{\mu} \cdot \frac{1-(K+1)\rho^K + K\rho^{K+1}}{(1-\rho)^2}.
 \end{aligned}$$

Подставляя значение p_0 и проводя очевидные действия, получим

$$T_q = \frac{(1-\rho)\rho}{\mu(1-\rho^{K+1})} \cdot \frac{[1-(K+1)\rho^K + K\rho^{K+1}]}{(1-\rho)^2} = \frac{\rho[1-(K+1)\rho^K + K\rho^{K+1}]}{\mu(1-\rho^{K+1})(1-\rho)}.$$

Таким образом, среднее время T_q пребывания одного требования в очереди равно

$$T_q = \frac{\rho[1-(K+1)\rho^K + K\rho^{K+1}]}{\mu(1-\rho^{K+1})(1-\rho)}. \quad (3.67)$$

С учетом того, что $\rho = \lambda / \mu$, можно связать среднее время T_q пребывания одного требования в очереди со средней длиной очереди N_q в виде

$$T_q = \frac{N_q}{\lambda}. \quad (3.68)$$

3.6. Система М/М/1//М с конечным числом источников нагрузки

Для рассмотренных выше систем предполагалось, что входной поток может поставлять бесконечное число требований или как говорят, является с бесконечным числом источников нагрузки. В обозначении

системы $M/M/1//M$ последняя справа буква M означает, что входной поток требований создается конечной группой возможных требований [4]. Система состоит из очереди и одного прибора обслуживания. Из всего M числа требований часть может находиться в очереди и одно требование в обслуживающем приборе. В частности, если в системе находятся k требований (очередь плюс прибор обслуживания), то $M-k$ требований находятся в числе поступающих. При этом, когда требование находится в группе поступающих, то момент времени его поступления в систему является случайной величиной, распределенной по экспоненциальному (показательному) закону (первая слева буква M в обозначении системы) со средним значением $1/\lambda$. Обслуживание в системе также происходит по экспоненциальному закону (вторая буква M в обозначении системы) со средним значением, равным $1/\mu$. Структура системы $M/M/1//M$ показана на рис. 3.9, на котором показан накопитель с $(k-1)$ требованиями и обслуживающий прибор с одним требованием. Входной поток для данного состояния системы состоит из $(M-k)$ требований.

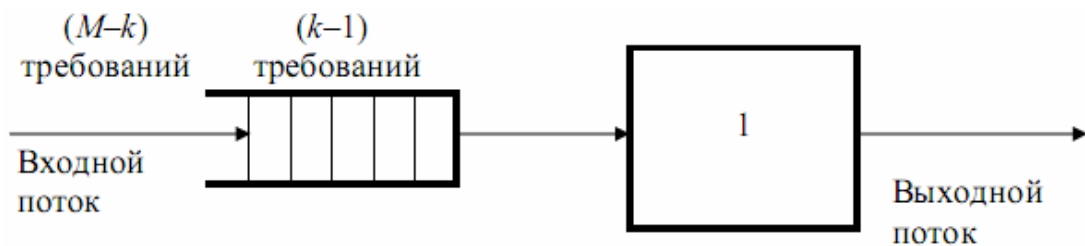


Рис. 3.9. Структура системы $M/M/1//M$

3.6.1. Вероятности состояний системы $M/M/1//M$

Для системы с M источниками нагрузки все поступающие требования действуют независимо друг от друга. В этом случае [4], если в системе находятся k требований, то $M-k$ требований находятся в числе

поступающих (см. рис.3.9), и тогда общая интенсивность поступления требований будет равна $\lambda(M-k)$. Очевидно, что с изменением числа требований, находящихся в системе, происходит изменение общей интенсивности поступающих требований. Система $M/M/1//M$ в определенном смысле является саморегулируемой. Для моделирования такой системы может быть использован процесс размножения и гибели со следующими параметрами:

$$\lambda_k = \begin{cases} \lambda(M-k), & 0 \leq k \leq M; \\ 0, & k > M; \end{cases} \quad (3.69)$$

$$\mu_k = \mu, \quad k = 1, 2, 3, \dots$$

С принятыми значениями интенсивностей входного потока и обслуживания может быть построен размеченный граф состояний системы $M/M/1//M$. Он показан на рис. 3.10.

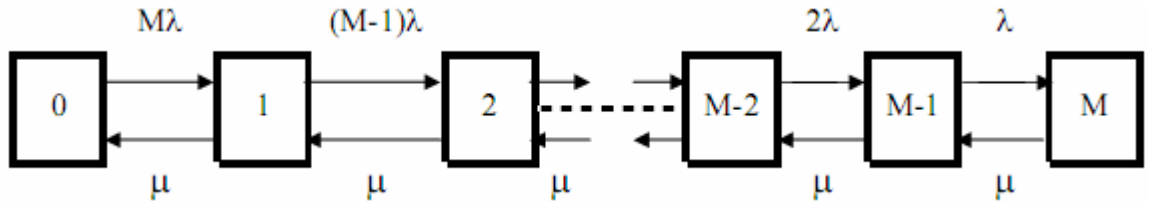


Рис. 3.10. Граф состояний системы $M/M/1//M$

В стационарном режиме вероятности состояний находим по общей формуле (2.46) процесса размножения и гибели с учетом заданных интенсивностей (3.69):

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda(M-i)}{\mu}, \quad 0 \leq k \leq M. \quad (3.70)$$

Выразим произведение (3.70):

$$\begin{aligned} p_k &= p_0 \prod_{i=0}^{k-1} \frac{\lambda(M-i)}{\mu} = p_0 \frac{\lambda M}{\mu} \cdot \frac{\lambda(M-1)}{\mu} \cdot \frac{\lambda(M-2)}{\mu} \cdot \dots \cdot \frac{\lambda(M-k+1)}{\mu} = \\ &= p_0 \left(\frac{\lambda}{\mu} \right)^k M \cdot (M-1) \cdot (M-2) \cdot \dots \cdot (M-k+1). \end{aligned}$$

Получим удобную компактную запись, используя факториалы и с учетом того, что $\lambda/\mu = \rho$:

$$p_k = p_0 \rho^k \frac{M!}{(M-k)!}, \quad 0 \leq k \leq M. \quad (3.71)$$

Определим вероятность p_0 из нормировочного условия:

$$\sum_{k=0}^M p_k = 1, \quad \sum_{k=0}^M p_0 \rho^k \frac{M!}{(M-k)!} = 1, \quad p_0 \sum_{k=0}^M \rho^k \frac{M!}{(M-k)!} = 1, \quad p_0 = \frac{1}{\sum_{k=0}^M \rho^k \frac{M!}{(M-k)!}}.$$

В другом виде

$$p_0 = \left[\sum_{k=0}^M \rho^k \frac{M!}{(M-k)!} \right]^{-1}. \quad (3.72)$$

3.6.2. Операционные характеристики системы М/М/1//М

Определим вероятность $P_{отк}$ отказа по формальному признаку. Требование, поступающее в систему, получает отказ в том случае, когда все места в очереди заняты ($M-1$) и один прибор обслуживает. Это означает, что в системе находится M требований. Поэтому с учетом (3.71), (3.72) будем иметь

$$P_{отк} = p_M = p_0 \rho^M \frac{M!}{(M-M)!} = \frac{\rho^M M!}{\left[\sum_{k=0}^M \frac{M!}{(M-k)!} \right]}. \quad (3.73)$$

Относительная пропускная способность Q_M определяет долю обслуженных требований. Вероятность $P_{обс}$ обслуженных требований и вероятность $P_{отк}$ не обслуженных требований образуют полную группу событий, т. е.

$$P_{обс} + P_{отк} = 1, \quad P_{обс} = 1 - P_{отк}.$$

Поэтому

$$Q_M = P_{обс} = 1 - P_{отк}. \quad (3.74)$$

Абсолютная пропускная способность A_M — это среднее число требований, поступающих в прибор обслуживания в единицу времени:

$$A_M = \lambda Q_M = \lambda(1 - P_{отк}). \quad (3.75)$$

Определим среднее \bar{N}_M число требований в системе М/М/1//М способом, который использовался в [8, 13]. В соответствии с размеченным графом состояний (см. рис.3.10) систему дифференциальных уравнений для вероятностей состояний $P_k(t)$ имеет вид, $k = \overline{0, M}$:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -M\lambda P_0(t) + \mu P_1(t), \\ \frac{dP_k(t)}{dt} &= -[(M-k)\lambda + \mu]P_k(t) + (M-k+1)\lambda P_{k-1}(t) + \mu P_{k+1}(t), \quad k = \overline{1, M-1}, \\ \frac{dP_M(t)}{dt} &= \lambda P_{M-1}(t) - \mu P_M. \end{aligned} \quad (3.76)$$

Второе уравнение в системе (3.76) справедливо для $k = 1, 2, \dots, M-1$. К системе (3.76) добавим естественные начальные условия:

$$P_0(0) = 1; \quad P_k(0) = 0, \quad k = \overline{1, M}. \quad (3.77)$$

Нормировочное условие для системы (3.76) имеет вид

$$\sum_{k=0}^M P_k(t) = 1. \quad (3.78)$$

Левые и правые части дифференциальных уравнений (3.76) в соответствии с номером уравнения умножим на k ($k=0,1,2,\dots, M$) и произведем их суммирование почленно:

$$\begin{aligned} \sum_{k=0}^M k \frac{dP_k(t)}{dt} = & 0(-M\lambda P_0(t) + \mu P_1(t)) + \\ & + \sum_{k=1}^{M-1} k \{ -[(M-k)\lambda + \mu]P_k(t) + (M-k+1)\lambda P_{k-1}(t) + \mu P_{k+1}(t) \} + \\ & + M[\lambda P_{M-1}(t) - \mu P_M] \end{aligned} \quad (3.79)$$

Очевидно, что уравнение (3.79) можно переписать в виде

$$\begin{aligned} \sum_{k=0}^M k \frac{dP_k(t)}{dt} = & \sum_{k=1}^{M-1} k \{ -[(M-k)\lambda + \mu]P_k(t) + (M-k+1)\lambda P_{k-1}(t) + \mu P_{k+1}(t) \} + \\ & + M[\lambda P_{M-1}(t) - \mu P_M] \end{aligned}$$

В левой части полученного уравнения поменяем местами операции суммирования и дифференцирования, а в правой части после почленного суммирования выполним преобразования для уравнивания показателей индексов:

$$\begin{aligned} \frac{d}{dt} \sum_{k=0}^M k P_k(t) = & -\lambda \sum_{k=0}^{M-1} k(M-k) P_k(t) + \lambda \sum_{k=0}^{M-1} (k+1)(M-k) P_k(t) - \\ & - \mu \sum_{k=1}^M k P_k(t) + \mu \sum_{k=1}^M (k-1) P_k(t). \end{aligned} \quad (3.80)$$

В правой части (3.80) имеем производную от математического ожидания, которое обозначим как $\bar{N}(t)$, в левой части раскроем скобки и приведем подобные члены:

$$\frac{d\bar{N}(t)}{dt} = \sum_{k=0}^{M-1} (M-k)\lambda P_k(t) - \sum_{k=1}^M \mu P_k(t).$$

Выполним еще преобразования для уравнивания индексов сумм:

$$\frac{d\bar{N}(t)}{dt} = \lambda M P_0(t) + \sum_{k=1}^{M-1} (M-k)\lambda P_k(t) - \sum_{k=1}^{M-1} \mu P_k(t) - \mu P_M(t).$$

Объединим две суммы с одинаковыми индексами суммирования, получим:

$$\frac{d\bar{N}(t)}{dt} = \lambda M P_0(t) + \sum_{k=1}^{M-1} [(M-k)\lambda - \mu] P_k(t) - \mu P_M(t).$$

Сумму в правой части равенства представим в виде следующих трех сумм:

$$\frac{d\bar{N}(t)}{dt} = \lambda M P_0(t) + \lambda M \sum_{k=1}^{M-1} P_k(t) - \lambda \sum_{k=1}^{M-1} k P_k(t) - \mu \sum_{k=1}^{M-1} P_k(t) - \mu P_M(t).$$

В правой части под знак суммы можно ввести слагаемые $\lambda M P_0(t)$ и $\mu P_M(t)$:

$$\frac{d\bar{N}(t)}{dt} = \lambda M \sum_{k=0}^{M-1} P_k(t) - \lambda \sum_{k=1}^{M-1} k P_k(t) - \mu \sum_{k=1}^M P_k(t). \quad (3.81)$$

В каждую из сумм добавим члены и затем их вычтем, чтобы прийти к нормировочному условию и к математическому ожиданию:

$$\frac{d\bar{N}(t)}{dt} = \lambda M \sum_{k=0}^M P_k(t) - \lambda M P_M(t) - \lambda \sum_{k=0}^M k P_k(t) + \lambda M P_M(t) - \mu \sum_{k=0}^M P_k(t) + \mu P_0(t).$$

Приводя подобные члены и учитывая нормировочное условие (3.78), а также представляя вторую сумму как математическое ожидание, получим

$$\frac{d\bar{N}(t)}{dt} = (\lambda M - \mu) - \lambda \bar{N}(t) + \mu P_0(t). \quad (3.82)$$

Получили дифференциальное уравнение относительно искомого математического ожидания числа требований в системе М/М/1//М. Решение уравнения (3.82) должно производиться одновременно с системой уравнений (3.76) с начальными условиями (3.77), к которым должно быть добавлено начальное условие относительно математического ожидания. Из условий (3.77) следует, что в начальный момент времени математическое ожидание как среднее количество требований равно нулю, т. е.

$$P_0(0) = 1, \Rightarrow \bar{N}(0) = 0. \quad (3.83)$$

Систему уравнений (3.76) с уравнением (3.82) в общем случае можно решить численными методами, что при современном уровне развития

методов решения дифференциальных уравнений не составляет особых трудностей, если учесть широкие возможности современных компьютерных вычислительных пакетов (MATLAB, MATHCAD, MATHEMATICA).

Для систем, описываемых процессами размножения и гибели, при соответствующих параметрах процесса существует стационарный режим. Поэтому всегда можно в результате численного решения дифференциальных уравнений (3.76), (3.82) получить установившееся решение и, следовательно, определить математическое ожидание (среднее количество требований в системе) для стационарного режима системы M/M/1//M.

Рассмотрим *пример* по определению временной зависимости математического ожидания в системе M/M/1//M. Пусть для системы M/M/1//7 значения интенсивностей поступления и обслуживания требований равны

$$\lambda = 1,98, \mu = 6.086.$$

Получим дифференциальные уравнения вида (3.76), (3.82), описывающие вероятности состояний системы и математическое ожидание числа требований в системе, при заданных M, λ, μ . Начальные условия принимаем естественные, т. е.

$$P_0(0) = 1; \quad P_k(0) = 0, \quad k = \overline{1, M}, \quad \bar{N}_M(0) = 0.$$

Дифференциальные уравнения представим в виде (3.76), не производя вычислений коэффициентов:

$$\begin{aligned}
 \frac{dP_0(t)}{dt} &= -M \lambda P_0(t) + \mu P_1(t), \\
 \frac{dP_1(t)}{dt} &= (M - 1 + 1) \lambda P_0(t) - [(M - 1) \lambda + \mu] P_1(t) + \mu P_2(t), \\
 \frac{dP_2(t)}{dt} &= (M - 2 + 1) \lambda P_1(t) - [(M - 2) \lambda + \mu] P_2(t) + \mu P_3(t), \\
 \frac{dP_3(t)}{dt} &= (M - 3 + 1) \lambda P_2(t) - [(M - 3) \lambda + \mu] P_3(t) + \mu P_4(t), \\
 \frac{dP_4(t)}{dt} &= (M - 4 + 1) \lambda P_3(t) - [(M - 4) \lambda + \mu] P_4(t) + \mu P_5(t), \\
 \frac{dP_5(t)}{dt} &= (M - 5 + 1) \lambda P_4(t) - [(M - 5) \lambda + \mu] P_5(t) + \mu P_6(t), \\
 \frac{dP_6(t)}{dt} &= (M - 6 + 1) \lambda P_5(t) - [(M - 6) \lambda + \mu] P_6(t) + \mu P_7(t), \\
 \frac{dP_7(t)}{dt} &= \lambda P_6(t) - \mu P_7(t), \\
 \frac{d\bar{N}(t)}{dt} &= \mu P_0(t) - \lambda \bar{N}(t) + (M \lambda - \mu).
 \end{aligned} \tag{3.84}$$

Систему уравнений (3.84) можно представить в векторной форме

$$\frac{dX}{dt} = AX + B, \tag{3.85}$$

где $X = [P_0(t), P_1(t), \dots, P_7(t), \bar{N}_M(t)]^T$, $B = [0, 0, 0, 0, 0, 0, 0, 0, (M \lambda - \mu)]^T$,

A — матрица размера 9×9 коэффициентов при переменных системы (3.84).

Решение примера выполним в системе MATLAB.

Программный код решения примера:

```
function MM1
clc, close all
%% Параметры системы
L = 1.98;
M = 6.086;
MM = 7;
```

```
A = [-MM*L, M, zeros(1,7);
      (MM-1+1)*L, -((MM-1)*L+M), M, zeros(1,6);
      0, (MM-2+1)*L, -((MM-2)*L+M), M, zeros(1,5);
      0, 0, (MM-3+1)*L, -((MM-3)*L+M), M, zeros(1,4);
      zeros(1,3), (MM-4+1)*L, -((MM-4)*L+M), M, zeros(1,3);
      zeros(1,4), (MM-5+1)*L, -((MM-5)*L+M), M, zeros(1,2);
      zeros(1,5), (MM-6+1)*L, -((MM-6)*L+M), M, 0;
      zeros(1,6), L, -M, 0;
      M, zeros(1,7), -L];
B = [zeros(length(A)-1,1);(MM*L-M)];

%% Вектор начальных условий
P0 = [1; zeros(length(A)-1,1)];
%% Интервал интегрирования
T = [0, 6];
[t,P] = ode45(@adif2,T,P0, [], A,B);

Prob = P(end,1:end-1);
fprintf('\n Вероятности: \n')
fprintf('\t%0.4f\n', Prob)
%% Количество требований в установившемся режиме
N = sum([0*P(end,1),1*P(end,2),2*P(end,3),3*P(end,4),...
4*P(end,5),5*P(end,6),6*P(end,7),7*P(end,8)])
disp(' Другой способ контроля:')
NN = sum(linspace(0,7,8).*P(end,1:end-1))

%% Диаграммы
figure(1),
```

```
plot(t,P(:,1:end-1),'linew',1.5),
grid on,
xlabel('\bf\it - - - - - t - - - - - '),
ylabel('\bf Вероятность'),
set(gcf,'color',[0.7,0.7,0.7]), ylim([0,1.05])

figure(2),
plot(t,P(:,end),'linew',2),
grid on,
xlabel('\bf\it - - - - - t - - - - - ')
ylabel('\bf\it N_{cp}\rm\bf(\itt\rm\bf)'),
ylim([0, max(P(:,end)) + max(P(:,end))/20])
set(gcf,'color',[0.7,0.7,0.7])

% М-файл описания правых частей уравнений
function f = adif2(t,P,A,B);
f = A*P + B;
```

В результате выполнения программы выводятся установившиеся значения вероятностей, производится расчет математического ожидания по формуле математического ожидания случайных дискретных величин (в программе переменные N и NN), строятся графики изменения во времени вероятностей состояний (рис.3.11) и график (рис. 3.12) изменения во времени математического ожидания системы М/М/1//7:

Вероятности:

0.0241

0.0549

0.1072

0.1744

0.2269

0.2215

0.1441

0.0469

$N =$

4.0004

Другой способ контроля:

$NN =$

4.0004

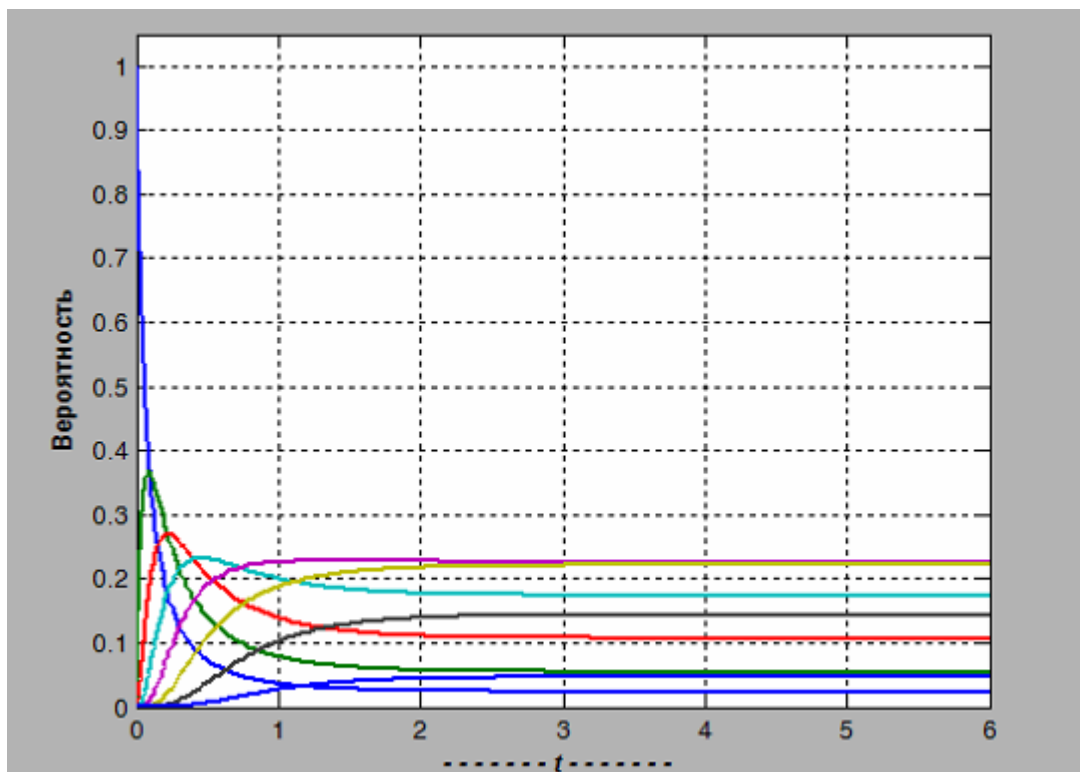


Рис. 3.11. Зависимость вероятностей состояний системы М/М/1//7 от времени

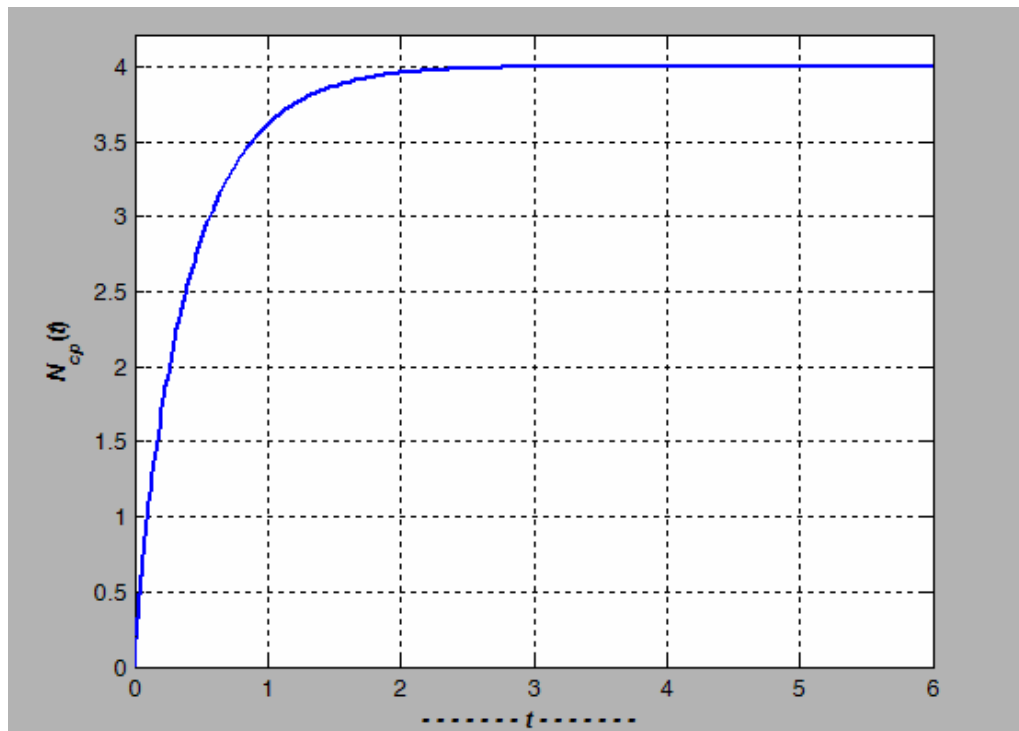


Рис. 3.12. Зависимость среднего количества требований
в системе М/М/1//7 от времени

Для сравнения результатов произведем расчет стационарных вероятностей (P_{st}) по формулам (3.71), (3.72) и затем определим среднее ($N_{control}$) количество требований в системе. Для этого используем следующую программу:

```
clear,clc
L = 1.98;
M = 6.086;
MM = 7;
r = L/M;
s1 = 0;
for k = 0 : MM
    s1 = s1 + r^k/factorial(MM-k);
end
P0 = 1/(factorial(MM)*s1);
```



```
q = 0;  
for k = 1 : MM  
    Pk(k) = factorial(MM)*P0*r^k/factorial(MM-k);  
end  
Pst = [P0,Pk]'  
Ncontrol = sum(linspace(0,7,8).*Pst')
```

Результат выполнения программы:

```
Pst =  
    0.0241  
    0.0549  
    0.1072  
    0.1744  
    0.2269  
    0.2215  
    0.1441  
    0.0469  
Ncontrol =  
    4.0004
```

Как видно, результаты совпадают: расчеты по прямым формулам и расчеты через решение дифференциальных уравнений.

В системе М/М/1//7 с ограниченным числом источников нагрузки среднее количество требований, очевидно, зависит от интенсивности обслуживания. Следует ожидать, что чем выше значение интенсивности обслуживания, тем меньше среднее количество требований будет в системе, все требования будут обслужены. Для иллюстрации этого

утверждения, произведем расчеты среднего количества требований в системе М/М/1//7 в зависимости от величины интенсивности μ . Для этого видоизменим приведенную выше программу, которая будет иметь следующий вид:

```
function MM1M;  
  
%% Параметры системы М/М/1//7  
L = 1.98;  
MM = 7;  
  
%% Вектор начальных условий  
P0 = [1; zeros(length(A)-1,1)];  
  
%% Интервал интегрирования дифференциальных уравнений  
T = [0, 6];  
  
for in = 0.5 : 0.5 : 50 % in - ИНТЕНСИВНОСТЬ ОБСЛУЖИВАНИЯ  
A = [-MM*L, in, zeros(1,MM);  
      (MM-1+1)*L, -((MM-1)*L+in), in, zeros(1,6);  
      0, (MM-2+1)*L, -((MM-2)*L+in), in, zeros(1,5);  
      0, 0, (MM-3+1)*L, -((MM-3)*L+in), in, zeros(1,4);  
      zeros(1,3), (MM-4+1)*L, -((MM-4)*L+in), in, zeros(1,3);  
      zeros(1,4), (MM-5+1)*L, -((MM-5)*L+in), in, zeros(1,2);  
      zeros(1,5), (MM-6+1)*L, -((MM-6)*L+in), in, 0;  
      zeros(1,6), L, -in, 0;  
      in, zeros(1,MM), -L];  
  
B = [zeros(length(A)-1,1);(MM*L-in)];
```

```
[t,P] = ode45(@adif2,T,P0,[],A,B);

%% Диаграмма
plot(in,P(end,end),'k','markersize',10),
grid on,hold on
end

text(51,0,'\bf\it\fontsize{12} \mu')
ylabel('\bf\it N_{cp}\rm\bf(\it\fontsize{12}\mu\rm\bf\fontsize{10})'),
xlabel('\bf\it\fontsize{12}\mu\rm\bf\fontsize{10} - интенсивность  
обслуживания'),
set(gcf,'color', [0.7,0.7,0.7])

%% М-файл описания правых частей уравнений
function f = adif2(t,P,A,B);
f = A*P + B;
```

Выполненные графические построения по программе приведены на рис. 3.13.

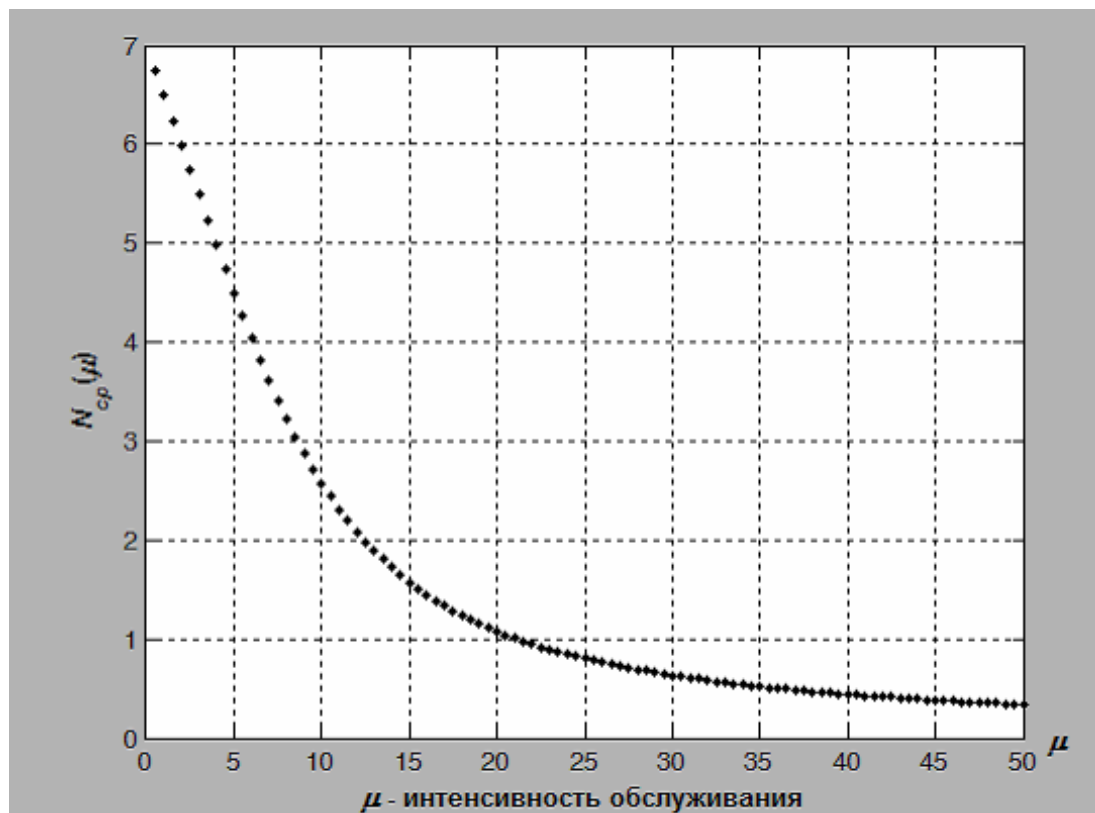


Рис. 3.13. Зависимость среднего числа требований от интенсивности обслуживания в системе М/М/1//7

Из рис. 3.13 видно, что характер изменения среднего числа требований от величины интенсивности обслуживания в системе типа М/М/1//М согласуется с логикой работы системы с ограниченным числом источников нагрузки. Можно сказать, что система типа М/М/1//М всегда является эргодической в смысле существования стационарного режима для вероятностей состояний, а, следовательно, и для значений среднего количества требований в системе.

3.7. Система М/М/1 с ограниченным временем ожидания

3.7.1. Вероятности состояний системы М/М/1 с ограниченным временем ожидания

Рассмотрим функционирование системы М/М/1, когда на длину очереди нет ограничений, но время пребывания требований в очереди конечно и случайно. В этом случае требования с ограниченным временем ожидания начала обслуживания называют "нетерпеливыми" требованиями. Структура системы массового обслуживания М/М/1 с "нетерпеливыми" требованиями показана на рис. 3.14.

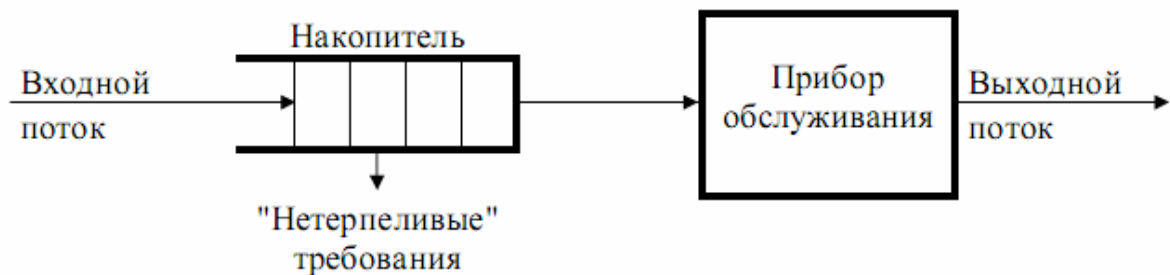


Рис. 3.14. Структура системы М/М/1 с "нетерпеливыми" требованиями

Условия работы системы:

1. Если в момент поступления требования прибор обслуживания свободен, то требование сразу начинает обслуживаться;
2. Если прибор обслуживания занят, то поступившее требование становится в очередь;
3. Если в момент освобождения прибора обслуживания имеется хотя бы одно требование в накопителе (в очереди), то первое из них по очереди сразу поступает на обслуживание;
4. В приборе обслуживания может находиться только одно требование;
5. Обслуживание не прерывается;

6. По окончании обслуживания требование покидает систему;
7. Очередь в накопителе упорядочена естественным образом, т. е. по дисциплине FIFO;
8. Требования могут покидать систему до начала обслуживания. Время ожидания в очереди "нетерпеливых" требований носит случайный характер;
9. Интервалы времени между моментом прихода и ухода "нетерпеливых" требований распределены по экспоненциальному закону;
10. Множество состояний системы бесконечно, т. е. нет ограничений на величину очереди.

Обозначим параметры системы:

λ — интенсивность поступления требований в систему,

μ — интенсивность обслуживания требований,

ν — интенсивность ухода "нетерпеливых" требований.

По сформулированным условиям работы системы принимаем, что время ожидания будет случайным и распределенным по экспоненциальному закону с плотностью $h(t)$ вида

$$h(t) = \nu e^{-\nu t}, \quad t > 0. \quad (3.86)$$

Анализ системы М/М/1 с ограниченным временем ожидания будем проводить по аналогии с анализом, рассматриваемым в [9, с.549] для многолинейных систем обслуживания.

Рассмотрим возможные состояния системы, которые будем обозначать через E с нижним индексом:

$$\left. \begin{aligned}
 E_0 & \text{ — прибор обслуживания не занят (очереди нет),} \\
 E_1 & \text{ — прибор обслуживания занят (очереди нет),} \\
 E_{1+1} & \text{ — прибор обслуживания занят, одно требование стоит} \\
 & \text{ в очереди,} \\
 E_{1+2} & \text{ — прибор обслуживания занят, два требования стоят} \\
 & \text{ в очереди,} \\
 & \dots\dots\dots \\
 E_{1+s} & \text{ — прибор обслуживания занят, } s \text{ требований стоят} \\
 & \text{ в очереди,} \\
 & \dots\dots\dots
 \end{aligned} \right\} (3.87)$$

По условиям работы системы число требований в очереди может быть бесконечно большим и поэтому рассматриваемая система имеет бесконечное множество состояний. Состояния E_0 и E_1 характеризуют систему без очереди. Переход в состояние E_1 может быть как из E_0 , так и из состояния E_2 , когда в очереди стоит одно требование. Получим дифференциальное уравнение для вероятности $P_1(t)$ состояния E_1 .

Возможны три способа перехода:

- 1) в момент времени t система уже была в состоянии E_1 , а за время Δt не вышла из него (требование не поступило, прибор обслуживания занят),
- 2) в момент t система была в состоянии E_0 , а за переходное время Δt перешла в состояние E_1 (поступило одно требование),
- 3) в момент t система была в состоянии E_2 , а за время Δt перешла в E_1 (либо освободился прибор обслуживания и стоящее в очереди требование заняло его, либо стоящее в очереди требование ушло в связи с окончанием времени ожидания обслуживания, ушло "нетерпеливое" требование).

Отообразим перечисленные возможности перехода в следующем виде:

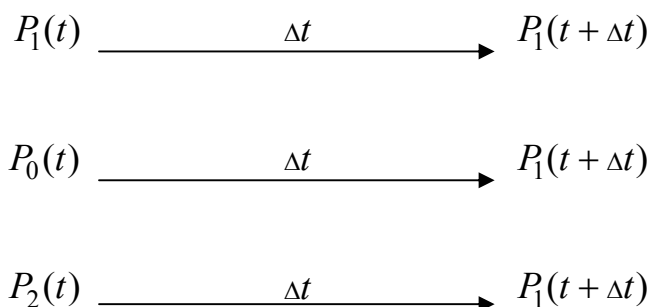


Рис. 3.15. Возможные переходы в состояние $P_1(t + \Delta t)$

Переход из одного состояния в другое происходит за время Δt , вероятность перехода за время Δt — это условная вероятность относительно безусловной вероятности на момент времени t . Отобразим взаимосвязь вероятностей в процессе перехода к $P_1(t + \Delta t)$ в следующем виде:

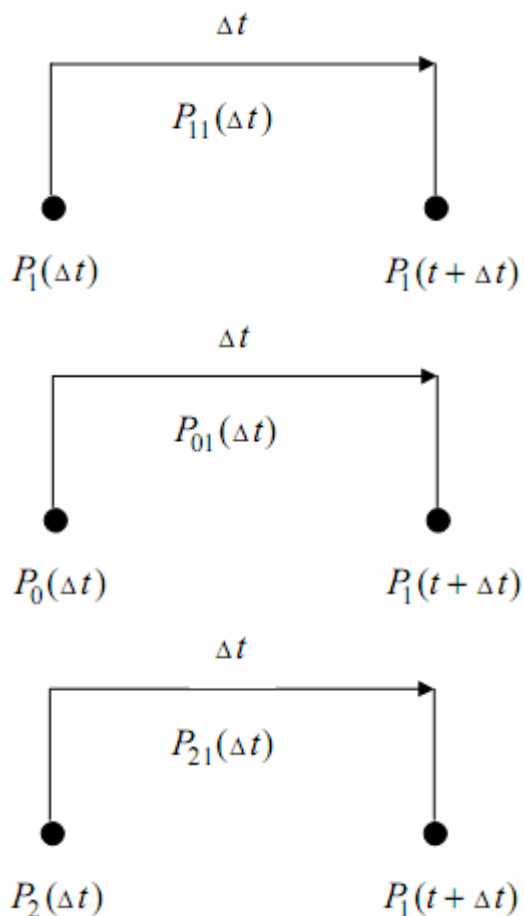


Рис. 3.16. Возможные переходы в состояние с вероятностью $P_1(t + \Delta t)$

Теперь воспользуемся формулой полной вероятности для определения вероятности $P_1(t + \Delta t)$:

$$P_1(t + \Delta t) = P_1(t)P_{11}(\Delta t) + P_0(t)P_{01}(\Delta t) + P_2(t)P_{21}(\Delta t), \quad (3.88)$$

где $P_{11}(\Delta t)$, $P_{01}(\Delta t)$, $P_{21}(\Delta t)$ — условные переходные вероятности.

Получим вероятности $P_{11}(\Delta t)$, $P_{01}(\Delta t)$, $P_{21}(\Delta t)$, принимая во внимание, что рассматриваемый поток простейший, т.е. стационарный, ординарный, с отсутствием последействия. Вероятность $P_{11}(\Delta t)$ перехода за время Δt из состояния E_1 в то же самое состояние E_1 означает, что новое требование не поступило и не произошло обслуживания. С учетом свойств простейшего потока имеем:

$$P_{11}(\Delta t) = 1 - \lambda \Delta t - \mu \Delta t + o(\Delta t), \quad (3.89)$$

где $o(\Delta t)$ — величина более высокого порядка малости, чем Δt .

Вероятность $P_{01}(\Delta t)$ перехода за время Δt из состояния E_0 в состояние E_1 означает, что переход происходит за счет поступления одного нового требования, поэтому

$$P_{01}(\Delta t) = \lambda \Delta t + o(\Delta t). \quad (3.90)$$

Вероятность $P_{21}(\Delta t)$ перехода за время Δt из состояния E_2 в состояние E_1 означает, что, либо произошло обслуживание одного требования, либо требование покинуло систему в связи с окончанием срока ожидания:

$$P_{21}(\Delta t) = \mu \Delta t + \nu \Delta t + o(\Delta t). \quad (3.91)$$

Значения вероятностей (3.89) — (3.91) подставим в (3.88):

$$P_1(t + \Delta t) = P_1(t)(1 - \lambda \Delta t - \mu \Delta t) + P_0(t)\lambda \Delta t + P_2(t)(\mu \Delta t + \nu \Delta t) + o(\Delta t),$$

где $o(\Delta t)$ — объединенная малая величина более высокого порядка малости, чем Δt .

Выполним очевидные преобразования:

$$P_1(t + \Delta t) - P_1(t) = -(\lambda + \mu)P_1(t)\Delta t + P_0(t)\lambda \Delta t + (\mu + \nu)P_2(t)\Delta t + o(\Delta t). \quad (3.92)$$

Обе части равенства (3.92) поделим на Δt и перейдем к пределу при $\Delta t \rightarrow 0$:

$$\lim_{\Delta t \rightarrow 0} \frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \left[\frac{-(\lambda + \mu)P_1(t)\Delta t}{\Delta t} + \frac{\lambda P_0(t)\Delta t}{\Delta t} + \frac{(\mu + \nu)P_2(t)\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t} \right].$$

Откуда получаем следующее дифференциальное уравнение:

$$\frac{dP_1(t)}{dt} = -(\lambda + \mu)P_1(t) + \lambda P_0(t) + (\mu + \nu)P_2(t). \quad (3.93)$$

Вычислим $P_{1+s}(t + \Delta t)$ при любом $s > 0$ — вероятность того, что в момент времени $t + \Delta t$ канал обслуживания будет занят, и ровно s требований будут стоять в очереди. Событие E_{1+s} , вероятность которого требуется вычислить, может осуществиться тремя способами:

- 1) в момент времени t система уже была в состоянии E_{1+s} , а за время Δt это состояние не изменилось, т. е. ни одно требование не поступило, прибор обслуживания не освободился, и ни одно из s стоящих в очереди требований не покинуло систему как "нетерпеливое" требование;
- 2) в момент времени t система была в состоянии E_{1+s-1} , а за время Δt перешла в состояние E_{1+s} , что означает поступление одного требования;
- 3) в момент времени t система была в состоянии E_{1+s+1} , а за время Δt перешла в состояние E_{1+s} : это либо прибор обслуживания освободится и одно требование из очереди займет его, либо одно из находящихся в очереди требований покинет систему как "нетерпеливое" требование.

По аналогии с формулой (3.88) будем иметь

$$P_{1+s}(t + \Delta t) = P_{1+s}(t)P_{1+s}(\Delta t) + P_{1+s-1}(t)P_{1+s-1}(\Delta t) + P_{1+s+1}(t)P_{1+s+1}(\Delta t), \quad (3.94)$$

где по условиям переходов в состояние E_{1+s} для простейшего потока переходные вероятности примем в виде:

$$P_{1+s}(\Delta t) = 1 - \lambda \Delta t - \mu \Delta t - s\nu \Delta t + o(\Delta t),$$

$$P_{1+s-1}(\Delta t) = \lambda \Delta t + o(\Delta t),$$

$$P_{1+s+1}(\Delta t) = \mu \Delta t + (s+1)\nu \Delta t + o(\Delta t).$$

Для первого равенства множитель s означает, что ни одно из s находящихся в очереди требований не покинуло систему. В третьем равенстве множитель $(s+1)$ означает, что любое из $(s+1)$ требований, находящихся в очереди, может покинуть систему. Проводя подстановки в (3.94) и очевидные преобразования, получим

$$P_{1+s}(t + \Delta t) = P_{1+s}(t)[1 - \lambda - \mu - s\nu]\Delta t + P_{1+s-1}(t)\lambda\Delta t + P_{1+s+1}(t)[\mu + (s+1)\nu]\Delta t + o(\Delta t).$$

Переносим в левую часть $P_{1+s}(t)$ и переходя к пределу, получим дифференциальное уравнение относительно вероятности $P_{1+s}(t)$ состояния E_{1+s} :

$$\frac{dP_{1+s}(t)}{dt} = -(\lambda + \mu + s\nu)P_{1+s}(t) + \lambda P_{1+s-1}(t) + [\mu + (s+1)\nu]P_{1+s+1}(t). \quad (3.95)$$

Теперь, по аналогии с (3.95) можно записать следующее дифференциальное уравнение относительно вероятности $P_{1+s+1}(t)$:

$$\frac{dP_{1+s+1}(t)}{dt} = -[\lambda + \mu + (s+1)\nu]P_{1+s+1}(t) + \lambda P_{1+s}(t) + [\mu + (s+2)\nu]P_{1+s+2}(t).$$

Таким образом, можно записать систему бесконечного числа дифференциальных уравнений для вероятностей состояний. Получим эту систему, принимая во внимание, что уравнение относительно вероятности отсутствия требований в системе будет иметь вид как в случае системы М/М/1 (первое уравнение в формуле с номером 3.2):

$$\begin{aligned}\frac{dP_0(t)}{dt} &= -\lambda P_0(t) + \mu P_1(t), \\ \frac{dP_1(t)}{dt} &= -(\lambda + \mu)P_1(t) + \lambda P_0(t) + (\mu + \nu)P_2(t), \\ \frac{dP_2(t)}{dt} &= -(\lambda + \mu + \nu)P_2(t) + \lambda P_1(t) + (\mu + 2\nu)P_3(t),\end{aligned}\tag{3.96}$$

$$\frac{dP_{1+s}(t)}{dt} = -(\lambda + \mu + s\nu)P_{1+s}(t) + \lambda P_{1+s-1}(t) + [\mu + (s+1)\nu]P_{1+s+1}(t),$$

В общем случае, стоящие в уравнениях системы (3.96), интенсивности λ, μ, ν могут быть переменными. Система (3.96) обладает той практической особенностью, что при увеличении очереди с числом s вероятности $P_{1+s}(t)$ становятся пренебрежимо малыми, и соответствующие уравнения могут быть отброшены.

В случае существования стационарного режима в системе М/М/1 с ограниченным временем ожидания в очереди система (3.96) будет иметь установившееся решение. Это решение может быть получено из соответствующей системы алгебраических уравнений, которые определяются, если в системе дифференциальных уравнений (3.96) все производные в левых частях положить равными нулю. Выпишем систему алгебраических уравнений, полагая, что в ней все искомые вероятности постоянные, которые обозначим строчными буквами вида p :

$$\begin{aligned}0 &= -\lambda p_0 + \mu p_1, \\ 0 &= -(\lambda + \mu)p_1 + \lambda p_0 + (\mu + \nu)p_2, \\ 0 &= -(\lambda + \mu + \nu)p_2 + \lambda p_1 + (\mu + 2\nu)p_3, \\ &\dots\dots\dots \\ 0 &= -(\lambda + \mu + s\nu)p_{1+s} + \lambda p_{1+s-1} + [\mu + (s+1)\nu]p_{1+s+1}, \\ &\dots\dots\dots\end{aligned}\tag{3.97}$$

К системе уравнений (3.97) следует присоединить нормировочное условие о выполнении условия полной группы событий:

$$\sum_{k=0}^{\infty} p_k = 1. \quad (3.98)$$

Начнем решать систему (3.97). Из первого уравнения найдем p_1 :

$$0 = -\lambda p_0 + \mu p_1, \Rightarrow p_1 = \frac{\lambda}{\mu} p_0. \quad (3.99)$$

Из второго уравнения системы (3.97) найдем p_2 с учетом значения p_1 :

$$0 = -(\lambda + \mu)p_1 + \lambda p_0 + (\mu + \nu)p_2, \Rightarrow p_2 = \frac{\lambda^2 p_0}{\mu(\mu + \nu)}.$$

Из третьего уравнения системы (3.97) найдем p_3 с учетом значения p_2 :

$$0 = -(\lambda + \mu + \nu)p_2 + \lambda p_1 + (\mu + 2\nu)p_3, \Rightarrow p_3 = \frac{\lambda^3 p_0}{\mu(\mu + \nu)(\mu + 2\nu)}.$$

Учитывая закономерность в определении вероятностей состояний, можно получить для любого $s \geq 1$, где s — длина очереди:

$$p_{1+s} = \frac{\lambda^{1+s} p_0}{\mu \prod_{i=1}^s (\mu + i\nu)}. \quad (3.100)$$

Вероятность p_0 , входящая в формулы (3.99), (3.100), определяется из нормировочного условия (3.98):

$$\sum_{k=0}^{\infty} p_k = 1, \quad p_0 + \sum_{k=1}^{\infty} p_k = 1, \quad p_0 + \frac{\lambda}{\mu} p_0 + \sum_{s=1}^{\infty} \frac{\lambda^{1+s} p_0}{\mu \prod_{i=1}^s (\mu + i\nu)} = 1,$$

$$p_0 \left[\frac{\lambda}{\mu} + \sum_{s=1}^{\infty} \frac{\lambda^{1+s}}{\mu \prod_{i=1}^s (\mu + i\nu)} \right] = 1, \Rightarrow p_0 = \frac{1}{\frac{\lambda}{\mu} + \sum_{s=1}^{\infty} \frac{\lambda^{1+s}}{\mu \prod_{i=1}^s (\mu + i\nu)}}. \quad (3.101)$$

Часто выражения (3.99), (3.100), (3.101) преобразуют при введении в них приведенных плотностей α и β :

$$\rho = \frac{\lambda}{\mu}, \quad \beta = \frac{\nu}{\mu}. \quad (3.102)$$

Параметры ρ и β выражают соответственно среднее число требований и среднее число уходов требования, стоящего в очереди, приходящиеся на среднее время обслуживания одного требования, т.к. $1/\mu$ — это и есть среднее время обслуживания одного требования в случае показательного закона обслуживания. С учетом обозначений (3.102) преобразуем формулы (3.99), (3.100), (3.101):

$$p_1 = p_0 \rho. \quad (3.103)$$

$$p_{1+s} = \frac{\lambda^{1+s} p_0}{\mu \prod_{i=1}^s (\mu + i\nu)} = \frac{\lambda^{1+s} p_0}{\mu^{1+s} \prod_{i=1}^s \left(1 + i \frac{\nu}{\mu}\right)} = \frac{\rho^{1+s} p_0}{\prod_{i=1}^s (1 + i\beta)}. \quad (3.104)$$

$$p_0 = \frac{1}{\rho + \sum_{s=1}^{\infty} \frac{\rho^{1+s}}{\prod_{i=1}^s (1 + i\beta)}}. \quad (3.105)$$

По формулам (3.103), (3.104), (3.105) рассчитываются стационарные вероятности состояний системы М/М/1 с ограниченным временем ожидания. Следует отметить, что вычисление бесконечной суммы, входящей в выражение (3.105), очевидно затруднительно. Но члены этой суммы быстро убывают, поэтому в практических расчетах можно ограничиться конечным числом членов исходной бесконечной суммы. При этом для грубой оценки ошибки, происходящей от отбрасывания всех членов сумм, начиная с r -го, можно воспользоваться формулой из [9, с.556], т. е.

$$\sum_{s=r}^{\infty} \frac{\rho^s}{\prod_{i=1}^s (1+i\beta)} < \frac{\left(\frac{\rho}{\beta}\right)^r}{r!} e^{\frac{\rho}{\beta}}. \quad (3.106)$$

Приведенная оценка (3.106) позволяет сделать заключение о возможности эргодического распределения вероятностей состояний. При $r \rightarrow \infty$ получаем следующую оценку:

$$\sum_{s=r}^{\infty} \frac{\rho^s}{\prod_{i=1}^s (1+i\beta)} < \frac{\left(\frac{\rho}{\beta}\right)^r}{r!} e^{\frac{\rho}{\beta}} \rightarrow 0. \quad (3.107)$$

Выполнение условия (3.107) означает, что эргодическое распределение (понимай стационарный режим системы) существует при любых $\rho, \beta > 0$ [5, с.93].

3.7.2. Операционные характеристики системы М/М/1 с ограниченным временем ожидания

Вычислим математическое ожидание N_s числа требований, находящихся в очереди:

$$N_s = \sum_{s=1}^{\infty} s p_{1+s} = \sum_{s=1}^{\infty} \frac{s \rho^{1+s} p_0}{\prod_{i=1}^s (1+i\beta)}, \quad (3.108)$$

где p_0 определяется по (3.105).

Вычислим вероятность P_n того, что требование покинет систему не обслуженным. Вероятность P_n определяется из следующих соображений: величина P_n — это отношение среднего числа требований, уходящих из очереди в единицу времени, к среднему числу требований, поступающих

в единицу времени [9, с. 555]. Таким образом, чтобы получить P_n , нужно N_s умножить на среднюю плотность уходов одного требования и разделить на среднюю плотность поступающих требований, т. е. умножить на такой коэффициент:

$$\frac{\frac{\nu}{\mu}}{\frac{\mu}{\lambda}} = \frac{\beta}{\rho} = \frac{\nu}{\lambda}. \quad (3.109)$$

С учетом (3.109) и (3.108) имеем

$$P_n = \frac{\beta}{\rho} \cdot \sum_{s=1}^{\infty} \frac{s \rho^{1+s} p_0}{\prod_{i=1}^s (1+i\beta)} = \sum_{s=1}^{\infty} \frac{\frac{\beta}{\rho} s \rho^{1+s} p_0}{\prod_{i=1}^s (1+i\beta)} = \sum_{s=1}^{\infty} \frac{\beta s \rho^s p_0}{\prod_{i=1}^s (1+i\beta)}. \quad (3.110)$$

Когда известна вероятность того, что требование будет не обслуженным, т.е. получит отказ, относительная пропускная способность Q системы вычисляется по формуле

$$Q = 1 - P_n = 1 - \sum_{s=1}^{\infty} \frac{\beta s \rho^s p_0}{\prod_{i=1}^s (1+i\beta)}. \quad (3.111)$$

Абсолютная пропускная способность A будет вычисляться по формуле

$$A = \lambda Q = \lambda(1 - P_n) = \lambda \left[1 - \sum_{s=1}^{\infty} \frac{\beta s \rho^s p_0}{\prod_{i=1}^s (1+i\beta)} \right]. \quad (3.112)$$

Непосредственное вычисление операционных характеристик по формулам (3.108), (3.110), (3.111), (3.112) затруднительно из-за присутствия в них бесконечных сумм. Но члены этих сумм быстро убывают [9], поэтому при практических расчетах можно ограничиться конечным числом членов сумм. При этом для грубой оценки ошибки,

происходящей от отбрасывания всех членов сумм, начиная с r -го, можно воспользоваться формулой из [9, с.556], т. е.

$$\sum_{s=1}^{\infty} \frac{s \rho^s}{\prod_{i=1}^s (1 + i\beta)} < \frac{\left(\frac{\rho}{\beta}\right)^r}{(r-1)!} e^{\frac{\rho}{\beta}}. \quad (3.113)$$

3.7.3. Пример моделирования системы М/М/1/8 с ограниченным временем ожидания

Проверим возможность существования стационарного режима для системы М/М/1/8 с ограниченным временем ожидания и ограниченным числом требований в системе. Пусть исходными параметрами системы будут: $\lambda = 2$, $\mu = 1.5$, $\nu = 1.2$. Число возможных состояний системы будет равно 9. Используя общую форму систем дифференциальных уравнений (3.96) и учитывая конечное число состояний 9, получим систему уравнений для рассматриваемого случая:

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t),$$

$$\frac{dP_1(t)}{dt} = -(\lambda + \mu)P_1(t) + \lambda P_0(t) + (\mu + \nu)P_2(t),$$

$$\frac{dP_2(t)}{dt} = -(\lambda + \mu + \nu)P_2(t) + \lambda P_1(t) + (\mu + 2\nu)P_3(t),$$

$$\frac{dP_3(t)}{dt} = -(\lambda + \mu + 2\nu)P_3(t) + \lambda P_2(t) + (\mu + 3\nu)P_4(t),$$

$$\frac{dP_4(t)}{dt} = -(\lambda + \mu + 3\nu)P_4(t) + \lambda P_3(t) + (\mu + 4\nu)P_5(t),$$

$$\frac{dP_5(t)}{dt} = -(\lambda + \mu + 4\nu)P_5(t) + \lambda P_4(t) + (\mu + 5\nu)P_6(t),$$

$$\frac{dP_6(t)}{dt} = -(\lambda + \mu + 5\nu)P_6(t) + \lambda P_5(t) + (\mu + 6\nu)P_7(t),$$

$$\frac{dP_7(t)}{dt} = -(\lambda + \mu + 6\nu)P_7(t) + \lambda P_6(t) + (\mu + 7\nu)P_8(t),$$

$$\frac{dP_8(t)}{dt} = \lambda P_7(t) - (\mu + 7\nu)P_8(t).$$

Для решения полученной однородной системы дифференциальных уравнений примем естественные начальные условия:

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k = \overline{1, 8}.$$

Проанализируем решение системы дифференциальных уравнений при изменении одного из параметров системы от 1 до 10, сохраняя два других неизменными. Моделирование поставленной задачи проведем в MATLAB по следующей программе, в которой исходные параметры системы: $L = \lambda$, $M = \mu$, $\nu = \nu$:

```
function MM18v;  
clc, close all  
M = 1.5;  
v = 1.2;  
A = zeros(9);  
%% Вектор начальных условий  
P0 = [1; zeros(length(A)-1, 1)];  
%% Интервал интегрирования дифференциальных уравнений  
T = [0, 15];  
  
%% L - интенсивность входного потока требований  
for L = 1 : 12  
  
A = [-L, M, 0, 0, 0, 0, 0, 0, 0;
```

```
L, -(L+M), (M+v), 0, 0, 0, 0, 0, 0;  
0, L, -(L+M+v), (M+2*v), 0, 0, 0, 0, 0;  
0, 0, L, -(L+M+2*v), (M+3*v), 0, 0, 0, 0;  
  
0, 0, 0, L, -(L+M+3*v), (M+4*v), 0, 0, 0;  
0, 0, 0, 0, L, -(L+M+4*v), (M+5*v), 0, 0;  
0, 0, 0, 0, 0, L, -(L+M+5*v), (M+6*v), 0;  
0, 0, 0, 0, 0, 0, L, -(L+M+6*v) (M+7*v);  
0, 0, 0, 0, 0, 0, 0, L, -(M+7*v)];
```

```
[t,P] = ode45(@puc317,T,P0,[], A);
```

```
figure(1),  
plot(t, P, 'linew', 1.1),  
hold on,  
grid on,
```

```
N = 0;  
for k = 2 : length(A)  
    N = N + (k-1)*P(:,k);  
end
```

```
figure(2),  
plot(t, N, 'linew', 1.5),  
hold on,  
grid on,  
end
```

```
figure(1),
text(-1,1.05,'\bf\it P\rm\bf(\itt\rm\bf)'),
ylim([0,1.05])
str = sprintf('%s (1 : %d); %s%g; %s%g',...
'\bf\fontsize{12}\lambda\fontsize{10} =', L, ...
'\fontsize{12}\mu\fontsize{10} =', M, ...
'v = ', v);
text(7.1,0.75, str),
xlabel('\bf\it - - - - - t - - - - - '),
ylabel('\bf Вероятность'),
set(gcf,'color',[0.7,0.7,0.7]),
hold off

figure(2),
text(-0.5,7.3,'\bf\it N\rm\bf(\itt\rm\bf)'),
text(15.1,0.01,'\bf\it t'),
xlabel(str);
set(gcf,'color', [0.7,0.7,0.7]),
hold off

% М-файл описания правых частей дифференциальных уравнений:
function f = puc317 (t,P, A);
f = A*P;
```

Диаграмма вероятностей состояний в функции времени при изменении интенсивности поступления требований от 1 до 12 показана на рис. 3.17. Зависимость среднего количества требований в функции времени показана на рис. 3.18.

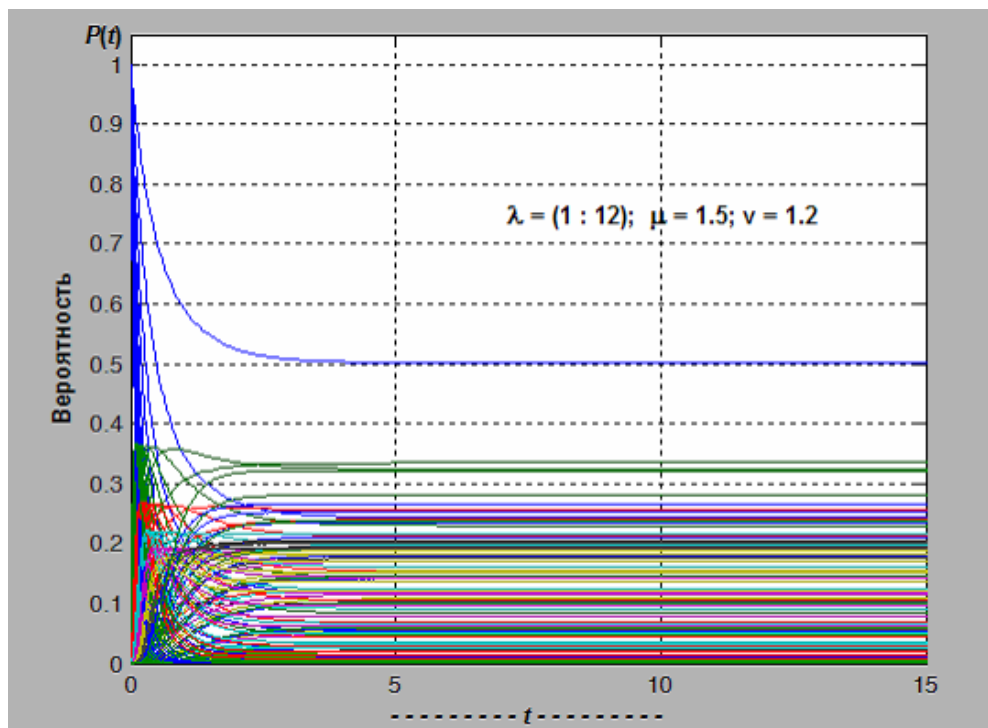


Рис. 3.17. Вероятности состояний системы М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности λ поступления требований

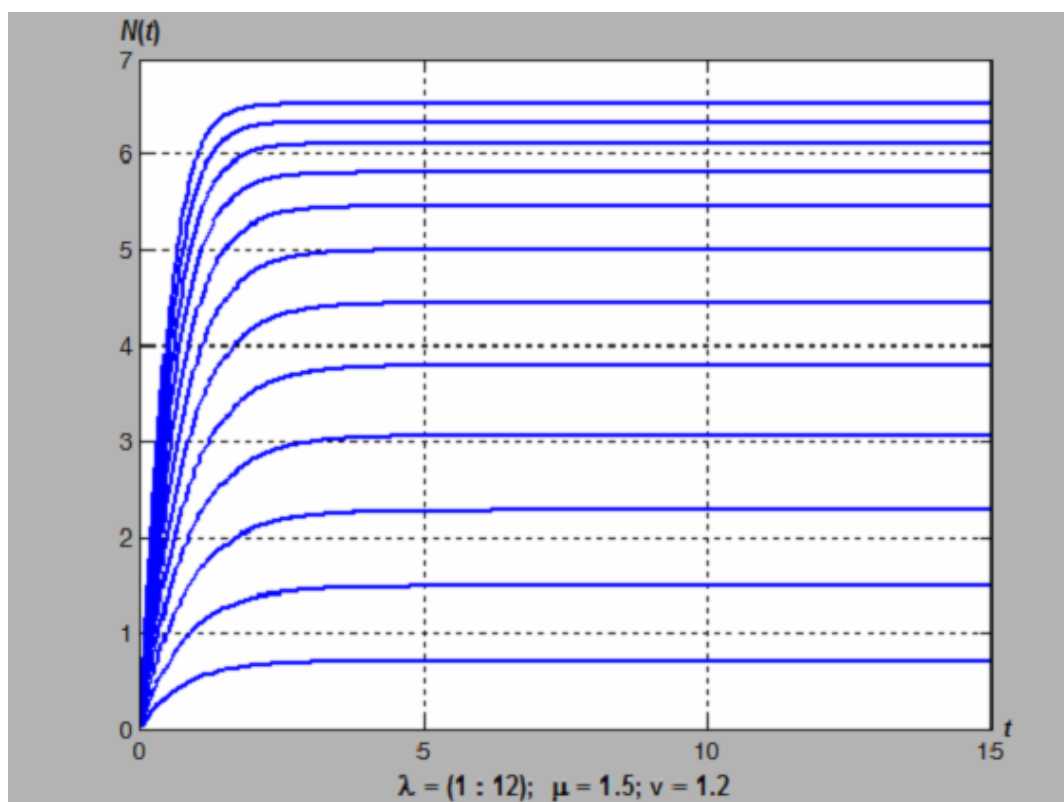


Рис. 3.18. Временная диаграмма среднего количества требований в системе при изменении интенсивности λ поступления требований

Как видно из рис. 3.17 исследуемая система обладает стационарным режимом. Вероятности состояний с течением времени стремятся к установившимся значениям. Тем самым подтверждается эргодичность системы $M/M/1/8$ с ограниченным временем ожидания при произвольных ее параметрах.

Приведенная программа позволяет рассчитывать и строить зависимость математического ожидания N количества требований в системе от времени t . График временной зависимости математического ожидания показан на рис. 3.18, на котором приведено семейство математических ожиданий при изменении интенсивности поступления требований в пределах от 1 до 12. Общим свойством поведения математического ожидания является стремление к установившемуся состоянию.

Для случая, когда циклически изменяется интенсивность обслуживания, программа принимает следующий вид:

```
function MM18
%% Очистка командного окна, закрытие графических окон
clc,close all

%% Параметры системы M/M/1/8 с нетерпеливыми требованиями
L = 2;
v = 1.2;

A = zeros(9);
%% Вектор начальных условий
P0 = [1;zeros(length(A)-1,1)];
%% Интервал интегрирования дифференциальных уравнений
T = [0, 15];
```

```
%% Цикл изменения интенсивности обслуживания
for M = 1 : 12

    A = [-L, M, 0, 0, 0, 0, 0, 0, 0;
        L, -(L+M), (M+v), 0, 0, 0, 0, 0, 0;
        0, L, -(L+M+v), (M+2*v), 0, 0, 0, 0, 0;
        0, 0, L, -(L+M+2*v), (M+3*v), 0, 0, 0, 0;
        0, 0, 0, L, -(L+M+3*v), (M+4*v), 0, 0, 0;
        0, 0, 0, 0, L, -(L+M+4*v), (M+5*v), 0, 0;
        0, 0, 0, 0, 0, L, -(L+M+5*v), (M+6*v), 0;
        0, 0, 0, 0, 0, 0, L, -(L+M+6*v), (M+7*v);
        0, 0, 0, 0, 0, 0, 0, L, -(M+7*v)];

    [t,P] = ode45(@pus319,T,P0,[],A);

    %% Диаграмма вероятностей
    figure(1),

    plot(t,P,'linew',1.1),
    hold on
    grid on,

    N = 0;
    for k = 2 : length(A)
        N = N + (k-1)*P(:,k);
    end

    %% Диаграмма изменения среднего количества требований
```

```
figure(2),
plot(t,N,'linew',1.5),
hold on,

end

figure(1),
grid on

text(-0.5,1.05,'\bf\it P\rm\bf(\itt\rm\bf)'),
text(15.1,0.01, '\bf\it t'),
ylim([0,1.01]),

str = sprintf('%s (1 : %d); %s%g; %s%g',...
'\bf\fontsize{12}\mu\fontsize{10} =', M, ...
'\fontsize{12}\lambda\fontsize{10} =', L, ...
'v = ', v);

text(6.9,0.95, str),
xlabel('\bf\it - - - - - t - - - - - '),
ylabel('\bf Вероятность')
set(gcf,'color',[0.7,0.7,0.7]),
hold off

figure(2),
grid on
text(-0.5,2.1, '\bf\it N\rm\bf(\itt\rm\bf)'),
text(15.1,0.01, '\bf\it t'),
```



```
xlabel(str),
set(gcf,'color',[0.7,0.7,0.7]),

%% М-файл описания правых частей дифференциальных уравнений
function f = puc319(t,P,A);
f = A*P;
```

Результат выполнения программы приведен на рис. 3.19, 3. 20, где показаны временные зависимости вероятностей состояний и среднего количества требований системы М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности обслуживания в пределах от 1 до 12.

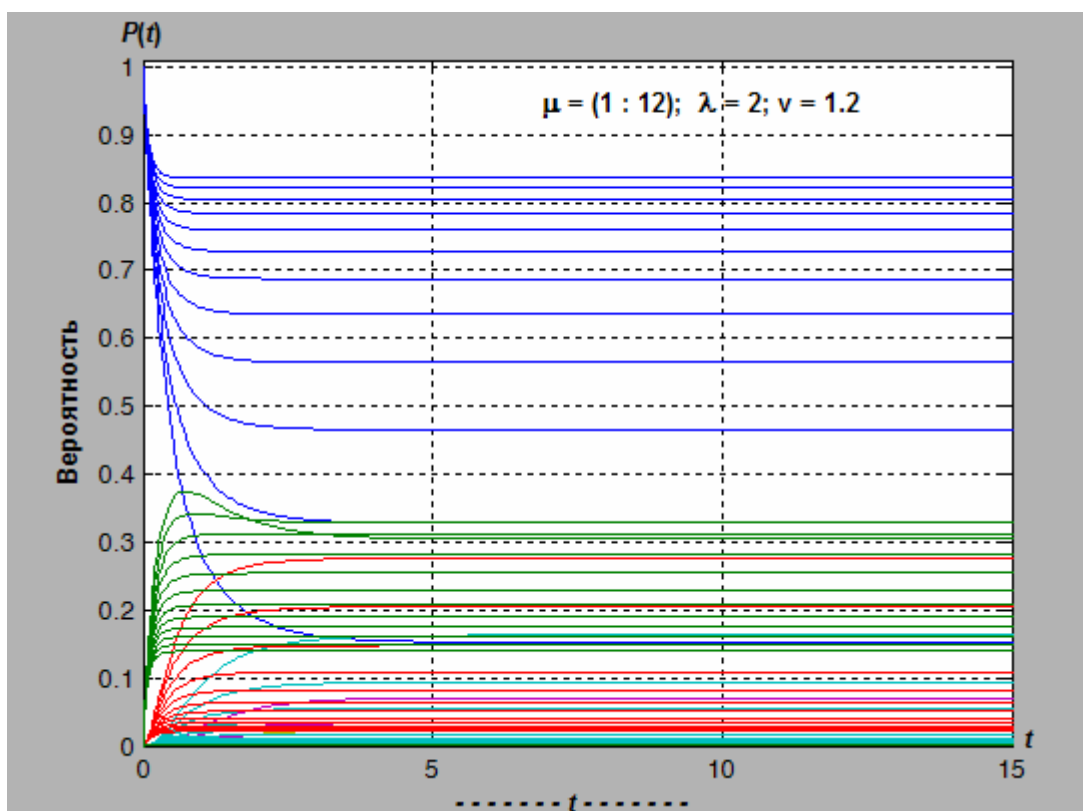


Рис. 3.19. Временная диаграмма вероятностей системы М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности обслуживания

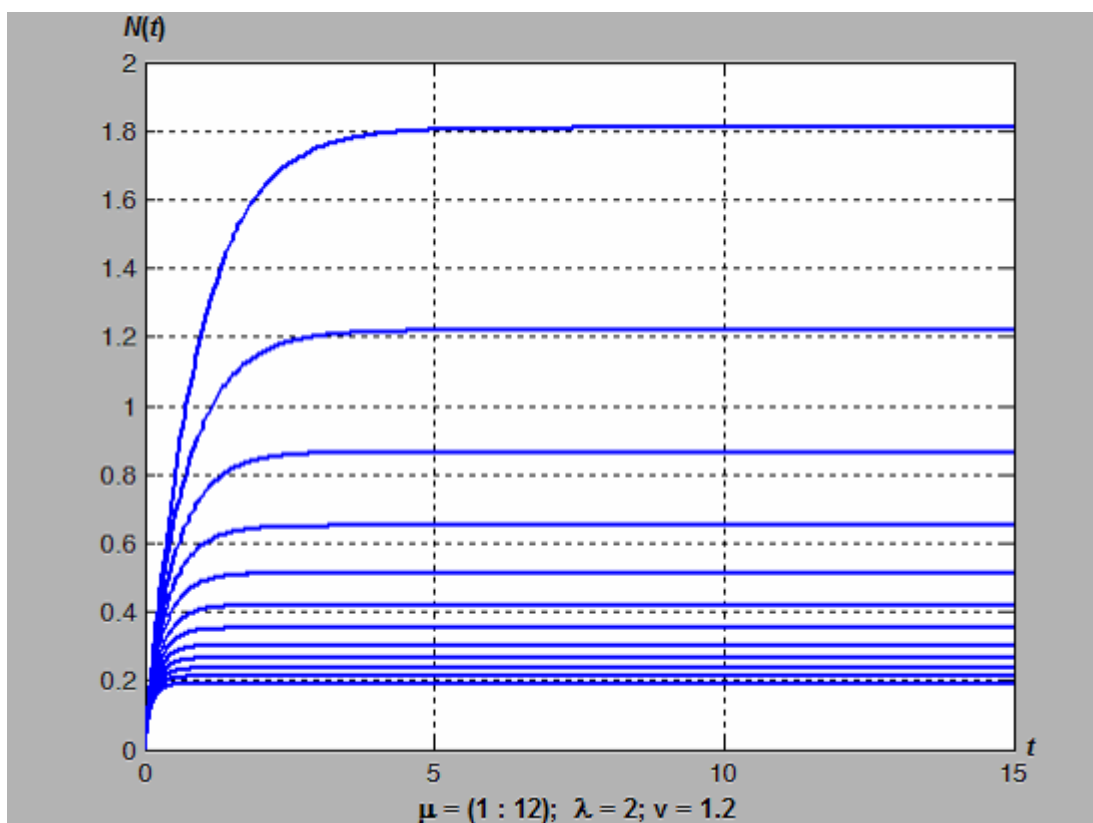


Рис. 3.20. Временная диаграмма среднего количества требований в системе М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности обслуживания

Для случая, когда циклически изменяется интенсивность ухода из очереди, программа принимает следующий вид:

```
function MM18
clc,close all
L = 2;
M = 1.5;

A = zeros(9);
P0 = [1;zeros(length(A)-1,1)];
T = [0, 15];
for v = 1 : 12
A = [-L, M, 0, 0, 0, 0, 0, 0, 0;
```

```
L, -(L+M), (M+v), 0, 0, 0, 0, 0, 0;  
0, L, -(L+M+v), (M+2*v), 0, 0, 0, 0, 0;  
0, 0, L, -(L+M+2*v), (M+3*v), 0, 0, 0, 0;  
0, 0, 0, L, -(L+M+3*v), (M+4*v), 0, 0, 0;  
0, 0, 0, 0, L, -(L+M+4*v), (M+5*v), 0, 0;  
0, 0, 0, 0, 0, L, -(L+M+5*v), (M+6*v), 0;  
0, 0, 0, 0, 0, 0, L, -(L+M+6*v), (M+7*v);  
0, 0, 0, 0, 0, 0, 0, L, -(M+7*v)];
```

```
[t,P] = ode45(@puc321,T,P0,[],A);
```

```
figure(1),
```

```
plot(t,P,'linew',1.1),
```

```
hold on
```

```
grid on,
```

```
N = 0;
```

```
for k = 2 : length(A)
```

```
    N = N + (k-1)*P(:,k);
```

```
end
```

```
figure(2),
```

```
plot(t,N,'linew',1.5),
```

```
hold on,
```

```
end
```

```
figure(1),
```

```
grid on
```

```
text(-0.5,1.05,'\bf{it P}\rm\bf{(\itt\rm\bf{)')},
```

```
text(15.1,0.01, '\bf\it t'),
ylim([0,1.01]),
str = sprintf('%s (1 : %d); %s%g; %s%g',...
'\bf v =', v, ...
'\fontsize{12}\lambda\fontsize{10} = ', L, ...
'\fontsize{12}\mu\fontsize{10} = ', M);
text(6.9,0.95, str),
xlabel('\bf\it - - - - - t - - - - - '),
ylabel('\bf Вероятность')
set(gcf,'color',[0.7,0.7,0.7]),
hold off

figure(2),
grid on
text(-0.5,1.9, '\bf\it N\rm\bf(\itt\rm\bf)'),
text(15.1,0.01, '\bf\it t'),

xlabel(str),
set(gcf,'color',[0.7,0.7,0.7]),

%% М-файл описания правых частей дифференциальных уравнений
function f = puc321(t,P,A);
f = A*P;
```

Временные зависимости вероятностей состояний и среднего количества требований в системе М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности ухода из очереди показаны на рис. 3.21, рис. 3.22.

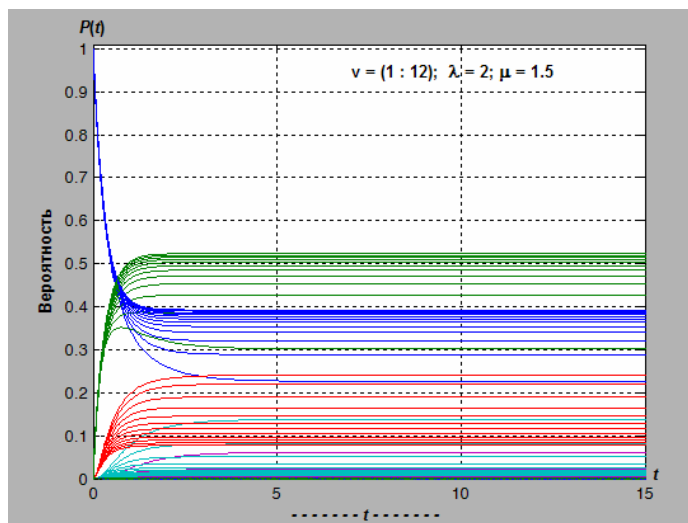


Рис. 3.21. Временная диаграмма вероятностей системы М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности ухода из очереди

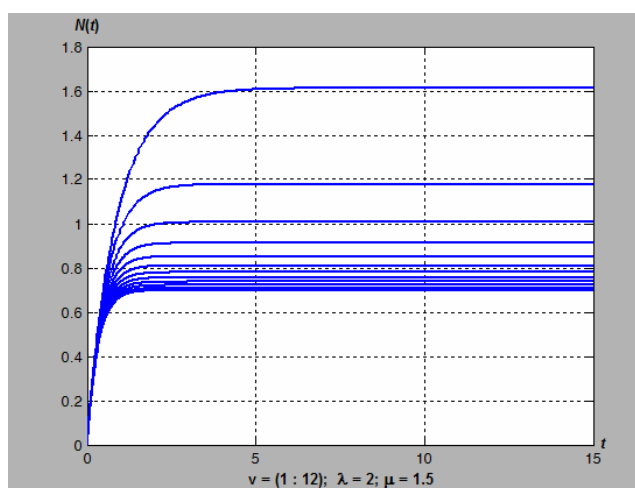


Рис. 3.22. Временная диаграмма среднего количества требований в системе М/М/1/8 с нетерпеливыми требованиями при изменении интенсивности ухода из очереди

Приведенные диаграммы подтверждают эргодичность системы, поскольку она обладает стационарным режимом, как для вероятностей состояний, так и для математического ожидания числа требований в системе.

ГЛАВА 4. МНОГОЛИНЕЙНЫЕ СИСТЕМЫ ОБСЛУЖИВАНИЯ

Как отмечалось во введении, многолинейные или многоканальные системы массового обслуживания представляют собой такие системы, у которых несколько приборов (каналов, линий, узлов) обслуживания соединены между собой параллельно, и поступающие требования направляются на свободный прибор, а после обслуживания покидают систему. Если в момент поступления требования свободны хотя бы два прибора, то требуется задать правило выбора того из них, на который поступит данное требование [5]. Будем предполагать, что каждый из свободных приборов может быть выбран с одной и той же вероятностью. В случае занятости всех приборов требования могут образовывать очередь, которая будет являться общей для всех приборов обслуживания. Модель многолинейной системы показана на рис. 4.1, где K_1, K_2, \dots, K_n — приборы обслуживания.

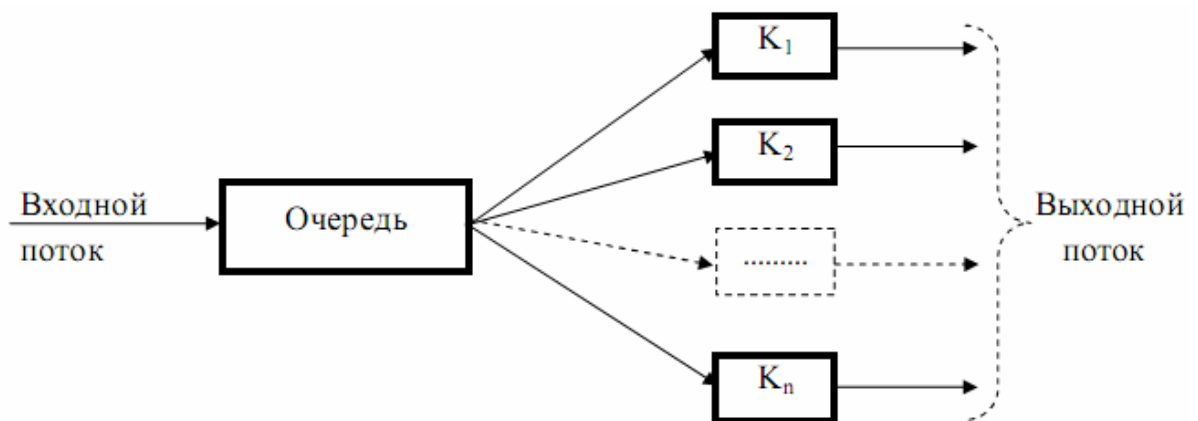


Рис. 4.1. Структура многолинейной системы обслуживания

При ограниченной длине очереди требования, застающие очередь заполненной, теряются и на обслуживание не поступают. Предполагается, что обслуживание требований осуществляется в порядке их поступления,

т. е. по дисциплине FIFO, обслуживание — по экспоненциальному закону, поступление требований — по пуассоновскому закону. В задачу анализа многолинейных систем входит определение вероятностей состояний и расчет операционных характеристик.

4.1. Система M/M/∞ с немедленным обслуживанием

Рассмотрим систему с бесконечным числом обслуживающих приборов, которые можно интерпретировать как наличие одного немедленно обслуживающего прибора, у которого интенсивность обслуживания растет линейно с ростом числа ожидающих требований [4]. В такой системе всегда найдется обслуживающий прибор, доступный вновь поступающему требованию. Параметры системы принимаются в виде:

$$\begin{aligned}\lambda_k &= \lambda = \text{const}, \quad k = 0, 1, 2, \dots; \\ \mu_k &= k\mu, \quad \mu = \text{const}, \quad k = 1, 2, 3, \dots\end{aligned}\tag{4.1}$$

Граф состояний системы M/M/∞ с параметрами (4.1) показан на рис. 4.2.

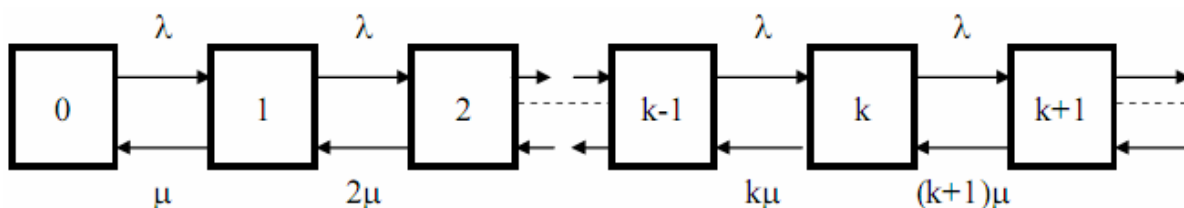


Рис. 4.2. Граф состояний системы M/M/∞

4.1.1. Вероятности состояний системы M/M/∞

Вероятностный процесс в системе M/M/∞ подчиняется процессу размножения и гибели с параметрами вида (4.1). Поэтому используем соотношения (2.46), (2.47) с учетом (4.1) для определения стационарных

(предельных) значений вероятностей состояния рассматриваемой системы обслуживания:

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} = p_0 \underbrace{\frac{\lambda}{\mu} \cdot \frac{\lambda}{2\mu} \cdot \dots \cdot \frac{\lambda}{k\mu}}_k = \frac{p_0}{k!} \left(\frac{\lambda}{\mu} \right)^k = \frac{\rho^k}{k!} p_0, \quad k = 0, 1, 2, \dots$$

$$p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu}} = \frac{1}{1 + \sum_{k=1}^{\infty} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k} = \frac{1}{1 + \sum_{k=1}^{\infty} \frac{\rho^k}{k!}} = \frac{1}{\sum_{k=0}^{\infty} \frac{\rho^k}{k!}} = e^{-\rho},$$

где

$$\rho = \frac{\lambda}{\mu}, \quad e^{\rho} = \sum_{k=0}^{\infty} \frac{\rho^k}{k!}.$$

Окончательное решение для вероятности p_k получаем в виде

$$p_k = \frac{\rho^k}{k!} e^{-\rho} = \frac{(\lambda/\mu)^k}{k!} e^{-\lambda/\mu}, \quad k = 0, 1, 2, \dots \quad (4.2)$$

Условие эргодичности для системы М/М/∞ задается очевидным неравенством

$$\frac{\lambda}{\mu} < \infty. \quad (4.3)$$

Неравенство (4.3) указывает на то, что отношение интенсивности поступления требований в систему к интенсивности обслуживания одним прибором может быть сколь угодно большим, но конечным числом. Какова бы ни была мала интенсивность обслуживания, в системе М/М/∞ всегда найдется прибор для обслуживания вновь поступившего требования.

4.1.2. Операционные характеристики системы М/М/∞

Вычислим среднее количество N требований в системе по формуле математического ожидания случайной дискретной величины:

$$N = \sum_{k=0}^{\infty} k p_k = \sum_{k=0}^{\infty} k \frac{\rho^k}{k!} e^{-\rho} = e^{-\rho} \sum_{k=1}^{\infty} \frac{\rho^k}{(k-1)!} = e^{-\rho} \rho \sum_{k=1}^{\infty} \frac{\rho^{(k-1)}}{(k-1)!}.$$

Переходя к новой переменной в последней сумме, будем иметь

$$N = e^{-\rho} \rho \sum_{q=0}^{\infty} \frac{\rho^q}{q!} = e^{-\rho} \rho e^{\rho} = \rho = \frac{\lambda}{\mu}, \quad (4.4)$$

где $q = k - 1$; $k = 1, q = 0$.

По формуле Литтла найдем среднее время T пребывания требования в системе:

$$N = \lambda T, \Rightarrow T = \frac{N}{\lambda} = \frac{\lambda / \mu}{\lambda} = \frac{1}{\mu}. \quad (4.5)$$

Для определения дисперсии D среднего числа требований в системе воспользуемся формулой

$$D = M(k^2) - [M(k)]^2, \quad (4.6)$$

где $M(k^2)$ — второй начальный момент, $M(k) = N$ согласно формуле (4.4).

Рассчитаем второй начальный момент:

$$\begin{aligned} M(k^2) &= \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=0}^{\infty} k^2 \frac{\rho^k}{k!} e^{-\rho} = \sum_{k=1}^{\infty} k \frac{\rho^k e^{-\rho}}{(k-1)!} = e^{-\rho} \sum_{k=1}^{\infty} k \frac{\rho^k}{(k-1)!} = \\ &= e^{-\rho} \rho \sum_{k=1}^{\infty} k \frac{\rho^{(k-1)}}{(k-1)!} = \left\{ \begin{array}{l} k-1 = z; \quad k=1, z=0; \\ k = z+1 \end{array} \right\} = e^{-\rho} \rho \sum_{z=0}^{\infty} (z+1) \frac{\rho^z}{z!} = \\ &= e^{-\rho} \rho \left[\sum_{z=0}^{\infty} z \frac{\rho^z}{z!} + \sum_{z=0}^{\infty} \frac{\rho^z}{z!} \right] = e^{-\rho} \rho \left[\sum_{z=1}^{\infty} \frac{\rho^z}{(z-1)!} + e^{\rho} \right] = \\ &= e^{-\rho} \rho \left[\rho \sum_{z=1}^{\infty} \frac{\rho^{(z-1)}}{(z-1)!} + e^{\rho} \right] = \left\{ \begin{array}{l} z-1 = q; \\ z=1, \quad q=0 \end{array} \right\} = e^{-\rho} \rho \left[\rho \sum_{q=0}^{\infty} \frac{\rho^q}{q!} + e^{\rho} \right] = \end{aligned}$$

$$= e^{-\rho} \rho \left[\rho e^{\rho} + e^{\rho} \right] = e^{-\rho} \rho^2 e^{\rho} + e^{-\rho} \rho e^{\rho} = \rho^2 + \rho.$$

Подставляя найденное значение $M(k^2)$ в (4.6) с учетом (4.4), получим

$$D = \rho^2 + \rho - \rho^2 = \rho = \frac{\lambda}{\mu}. \quad (4.7)$$

Очередь в системе $M/M/\infty$ не образуется, так как для вновь поступающего требования всегда найдется прибор обслуживания.

4.2. Система $M/M/m$ с m обслуживающими приборами

Системы типа $M/M/m$ — это системы с m обслуживающими приборами, работающими параллельно, и с неограниченным числом мест ожидания в очереди. Описание функционирования системы может быть сделано в терминах процесса размножения и гибели при следующих параметрах:

$$\begin{aligned} \lambda_k &= \lambda = \text{const}, \quad k = 0, 1, 2, \dots; \\ \mu_k &= \begin{cases} k\mu, & \mu = \text{const}, \quad 0 \leq k \leq m; \\ m\mu, & k \geq m. \end{cases} \end{aligned} \quad (4.8)$$

Размеченный граф состояний системы $M/M/m$ с параметрами (4.8) показан на рис. 4.3. Интенсивность обслуживания пропорциональна числу обслуживающих приборов до момента, когда состояние системы равно m . В состоянии 1 интенсивность обслуживания равна μ . В состоянии 2 интенсивность обслуживания равна 2μ , так как два канала одновременно могут обслужить два требования за одно и то же время, а это равносильно тому, что интенсивность обслуживания увеличивается в два раза. Когда в системе образуется очередь, то интенсивность обслуживания остается неизменной и равной $m\mu$, так как только m приборов могут обслужить требования.

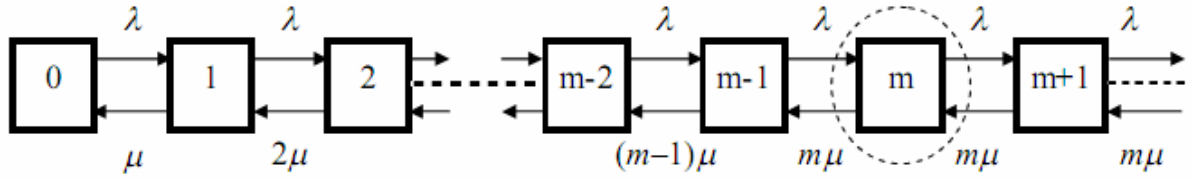


Рис. 4.3. Граф состояний системы М/М/т

4.2.1. Вероятности состояний системы М/М/т

Стационарные вероятности p_k (в предположении пока, что они существуют) состояний системы М/М/т будем определять в два этапа: 1-й этап, когда $k \leq m$, и 2-й этап, когда $k \geq m$.

В случае 1-го этапа применим общее выражение (2.46) с учетом (4.8):

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} = p_0 \underbrace{\frac{\lambda}{\mu} \cdot \frac{\lambda}{2\mu} \cdot \dots \cdot \frac{\lambda}{k\mu}}_{k \text{ раз}} = p_0 \frac{(\lambda/\mu)^k}{k!} = p_0 \frac{\rho^k}{k!}, \quad (4.9)$$

где $\rho = \lambda/\mu$.

В случае 2-го этапа, когда $k \geq m$, вывод решения для вероятностей состояния проведем с использованием графа состояний рис. 4.3. Рассмотрим состояние E_m , которое пунктирной линией обведено воображаемой границей. На основе входящих потоков через воображаемую границу и выходящих через эту границу составим следующее уравнение баланса для стационарного режима:

$$\lambda p_{m-1} + m\mu p_{m+1} = (\lambda + m\mu) p_m. \quad (4.10)$$

Разрешим уравнение (4.10) относительно p_{m+1} :

$$p_{m+1} = \frac{(\lambda + m\mu) p_m}{m\mu} - \frac{\lambda p_{m-1}}{m\mu}. \quad (4.11)$$

Значения p_m и p_{m-1} находим по формуле (4.9) и подставляем в (4.11):

$$\begin{aligned}
 p_{m+1} &= \frac{\lambda p_0 \rho^m}{m\mu m!} + \frac{p_0 \rho^m}{m!} - \frac{\lambda p_0 \rho^{(m-1)}}{m\mu(m-1)!} = \frac{p_0 \rho^m}{m!} + \frac{\lambda p_0 \rho^{(m-1)}}{m\mu(m-1)!} \left[\frac{\rho}{m} - 1 \right] = \\
 &= \frac{p_0 \rho^m}{m!} + \frac{p_0 \rho^m}{m(m-1)!} \left[\frac{\rho}{m} - 1 \right] = \frac{p_0 \rho^m}{m!} + \frac{p_0 \rho^m}{m!} \left[\frac{\rho}{m} - 1 \right] = \frac{p_0 \rho^m}{m!} \left(1 + \frac{\rho}{m} - 1 \right) = \frac{p_0 \rho^m \rho}{m! m} = \\
 &= p_0 \frac{\rho^{(m+1)}}{m! m}.
 \end{aligned}$$

Таким образом, имеем

$$p_{m+1} = p_0 \frac{\rho^{(m+1)}}{m! m}. \quad (4.12)$$

Рассмотрим следующее состояние E_{m+1} и составим уравнение баланса:

$$\lambda p_m + m\mu p_{m+2} = (\lambda + m\mu) p_{m+1}. \quad (4.13)$$

Уравнение (4.13) разрешим относительно p_{m+2} :

$$p_{m+2} = \frac{(\lambda + m\mu) p_{m+1}}{m\mu} - \frac{\lambda p_m}{m\mu}. \quad (4.14)$$

Подставим в (4.14) значения вероятностей p_{m+1} и p_m из (4.9), (4.12):

$$\begin{aligned}
 p_{m+2} &= \frac{(\lambda + m\mu) p_{m+1}}{m\mu} - \frac{\lambda p_m}{m\mu} = \frac{\lambda p_0 \rho^{(m+1)}}{m\mu m! m} + \frac{p_0 \rho^{(m+1)}}{m! m} - \frac{\lambda p_0 \rho^m}{m\mu m!} = \\
 p_{m+2} &= \frac{p_0 \rho \rho^{(m+1)}}{m! m^2} + \frac{p_0 \rho^{(m+1)}}{m! m} - \frac{p_0 \rho \rho^m}{m! m} = \frac{p_0 \rho^{(m+2)}}{m! m^2} + \frac{p_0 \rho^{(m+1)}}{m! m} - \frac{p_0 \rho^{(m+1)}}{m! m} = \\
 &= p_0 \frac{\rho^{(m+2)}}{m! m^2}.
 \end{aligned}$$

Получили выражение для p_{m+2} :

$$p_{m+2} = p_0 \frac{\rho^{(m+2)}}{m! m^2}. \quad (4.15)$$

Продолжая по аналогии дальнейшие вычисления, можно получить следующее выражение для вероятности состояния с номером $m+r$, где r — число, превышающее количество обслуживающих приборов:

$$p_{m+r} = p_0 \frac{\rho^{(m+r)}}{m! m^r}. \quad (4.16)$$

Преобразуем (4.16), принимая $k = m + r$:

$$p_k = p_0 \frac{\rho^k}{m! m^{k-m}}, \quad k \geq m. \quad (4.17)$$

Полученные соотношения (4.9), (4.17) для вероятностей состояний объединим и представим в следующем виде:

$$p_k = \begin{cases} p_0 \frac{\rho^k}{k!}, & k \leq m; \\ p_0 \frac{\rho^k}{m! m^{k-m}}, & k \geq m. \end{cases} \quad (4.18)$$

Учитывая, что в (4.18) $\rho = \lambda/\mu$, выполним следующие преобразования:

$$p_k = \begin{cases} p_0 \frac{\rho^k}{k!} = p_0 \frac{(m\lambda / m\mu)^k}{k!} = p_0 \frac{(mq)^k}{k!}, & k \leq m; \\ p_0 \frac{\rho^k}{m! m^{k-m}} = p_0 \frac{(\lambda / \mu)^k}{m! m^k m^{-m}} = p_0 \frac{(\lambda / m\mu)^k m^m}{m!} = p_0 \frac{q^k m^m}{m!}, & k \geq m, \end{cases}$$

где $q = \lambda/m\mu$.

Получаем расчетные формулы

$$p_k = \begin{cases} p_0 \frac{(mq)^k}{k!}, & k \leq m; \\ p_0 \frac{q^k m^m}{m!}, & k \geq m. \end{cases} \quad (4.19)$$

По (4.19) легко установить для системы М/М/т условие эргодичности (здесь условие существования стационарного режима). Каждая из дробей в (4.19) при любом k должна быть меньше единицы, что достигается при выполнении следующего неравенства:

$$q = \frac{\lambda}{m\mu} < 1. \quad (4.20)$$

Неравенство (4.20) определяет собой условие эргодичности для системы М/М/т с конечным числом t обслуживающих приборов и с неограниченным числом мест для ожидания.

Теперь определим вероятность p_0 — вероятность отсутствия требований в системе М/М/т. Для этого обратимся к нормировочному условию, учитывая (4.19):

$$\begin{aligned} \sum_{k=0}^{\infty} p_k = 1, & \Rightarrow \sum_{k=0}^{m-1} p_0 \frac{(mq)^k}{k!} + \sum_{k=m}^{\infty} p_0 \frac{q^k m^m}{m!} = p_0 + \sum_{k=1}^{m-1} p_0 \frac{(mq)^k}{k!} + \sum_{k=m}^{\infty} p_0 \frac{q^k m^m}{m!} = 1, \\ p_0 \left[1 + \sum_{k=1}^{m-1} \frac{(mq)^k}{k!} + \sum_{k=m}^{\infty} \frac{q^k m^m}{m!} \right] &= 1, \quad p_0 = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{1}{m!} \sum_{k=m}^{\infty} \frac{(\lambda / m\mu)^k}{m^{-m}} \right]^{-1} = \\ &= \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{1}{m!} \sum_{k=m}^{\infty} \frac{(\lambda / \mu)^k}{m^k m^{-m}} \right]^{-1} = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{m^m}{m!} \sum_{k=m}^{\infty} \left(\frac{\lambda}{m\mu} \right)^k \right]^{-1} = \\ &= \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{m^m}{m!} \sum_{k=m}^{\infty} q^k \right]^{-1} = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{m^m q^m}{m!(1-q)} \right]^{-1} = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!(1-q)} \right]^{-1}, \end{aligned}$$

где была использована формула бесконечно убывающей геометрической прогрессии с первым членом q^m и знаменателем q .

Таким образом, получили

$$p_0 = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!(1-q)} \right]^{-1}, \quad (4.21)$$

где

$$q = \frac{\lambda}{m\mu}.$$

Вероятность p_0 может интерпретироваться как вероятность, которая определяет среднюю долю времени простоя системы обслуживания.

4.3. Операционные характеристики системы М/М/м

4.3.1. Вероятность ожидания начала обслуживания. С-формула Эрланга

Рассчитаем вероятность $P_{ож}$ ожидания начала обслуживания или вероятность того, что вновь поступившее требование окажется в очереди. Очередь возникает в том случае, когда все приборы обслуживания окажутся занятыми. Следовательно, необходимо подсчитать вероятность присутствия в системе не менее m требований. Поскольку появления требований в системе являются событиями несовместными, то можно применить теорему сложения вероятностей. Тогда искомая характеристика $P_{ож}$ определяется в виде

$$P_{ож} = \sum_{k=m}^{\infty} p_k = \sum_{k=m}^{\infty} p_0 \frac{(mq)^k}{m! m^{k-m}} = p_0 \frac{m^m}{m!} \sum_{k=m}^{\infty} \frac{(mq)^k}{m^k} = p_0 \frac{m^m}{m!} \sum_{k=m}^{\infty} q^k.$$

Учитывая (4.20), т. е. то, что $q < 1$, применим формулу бесконечно убывающей геометрической прогрессии:

$$P_{ож} = p_0 \frac{m^m}{m!} \cdot \frac{q^m}{(1-q)} = p_0 \frac{(mq)^m}{m!(1-q)}. \quad (4.22)$$

С учетом (4.21) формула (4.22) приобретает вид

$$P_{ож} = \frac{\frac{(mq)^m}{m!(1-q)}}{\left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!(1-q)} \right]}. \quad (4.23)$$

Формулу (4.23) широко используют в телефонии; она определяет вероятность того, что поступающий на пучок m линий вызов (требование) не застанет ни одной свободной линии (обслуживающего прибора); рассматриваемая формула называется *С-формулой Эрланга* [4, с.121].

4.3.2. Средняя длина очереди в системе М/М/т

Рассчитаем среднее количество N_q находящихся в очереди требований. Для этого следует определить математическое ожидание количества требований, находящихся в очереди. Каждому требованию, находящемуся в очереди, соответствует вероятность состояния, которое есть число обслуживающих приборов плюс число требований в очереди. Расчет N_q произведем по формуле математического ожидания случайной дискретной величины (количества требований):

$$\begin{aligned} N_q &= \sum_{k=m}^{\infty} (k-m) p_k = \left\{ \begin{array}{l} k-m=i, k=m+i \\ k=m, \quad i=0 \end{array} \right\} = \sum_{i=0}^{\infty} i p_{m+i} = \sum_{i=0}^{\infty} i p_0 \frac{q^{m+i} m^m}{m!} = \\ &= p_0 \frac{m^m q^m}{m!} \sum_{i=0}^{\infty} i q^i = p_0 \frac{m^m q^m}{m!} q \frac{d}{dq} \sum_{i=0}^{\infty} q^i = p_0 \frac{m^m q^m}{m!} q \frac{d}{dq} \left(\frac{1}{1-q} \right) = \\ &= p_0 \frac{m^m q^m q}{m!} \cdot \left[\frac{1}{(1-q)^2} \right] = p_0 \frac{(mq)^m q}{m!} \left[\frac{1}{(1-q)^2} \right] = p_0 \frac{\left(\frac{m\lambda}{m\mu} \right)^m \frac{\lambda}{m\mu}}{m! \left(1 - \frac{\lambda}{m\mu} \right)^2} = \\ &= p_0 \frac{(m\mu)^2 \left(\frac{\lambda}{\mu} \right)^m \frac{\lambda}{m\mu}}{m! (m\mu - \lambda)^2} = p_0 \frac{m\mu \rho^m \lambda}{m! \mu^2 (m - \rho)^2} = p_0 \frac{\rho^{m+1}}{(m-1)! (m - \rho)^2}, \end{aligned}$$

где

$\rho = \frac{\lambda}{\mu}$, m — количество приборов обслуживания системы М/М/т, p_0 — вероятность отсутствия требований в системе, определяется по формуле (4.21).

Получили следующее выражение для расчета среднего количества требований в очереди системы М/М/т:

$$N_q = p_0 \frac{\rho^{m+1}}{(m-1)!(m-\rho)^2}. \quad (4.24)$$

С учетом (4.21) и величины $q = \lambda/m\mu$ значение N_q приобретает вид

$$\begin{aligned} N_q &= \frac{\rho^{m+1}}{(m-1)!(m-\rho)^2 \left(\sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^m}{(m-1)!(m-\rho)} \right)} = \\ &= \frac{\rho^{m+1}}{\left[\sum_{k=0}^{m-1} \frac{(m-1)!(m-\rho)^2 \rho^k}{k!} + (m-\rho)\rho^m \right]}. \end{aligned} \quad (4.25)$$

4.3.3. Среднее время ожидания требования в очереди начала обслуживания

Расчет среднего времени ожидания T_q одного требования в очереди начала обслуживания произведем по формуле Литтла:

$$N_q = \lambda T_q, \Rightarrow T_q = \frac{N_q}{\lambda}.$$

С учетом (4.25) имеем:

$$\begin{aligned} T_q &= \frac{\rho^{m+1}}{\lambda \left[\sum_{k=0}^{m-1} \frac{(m-1)!(m-\rho)^2 \rho^k}{k!} + (m-\rho)\rho^m \right]} = \\ &= \frac{\rho^m}{\mu \left[\sum_{k=0}^{m-1} \frac{(m-1)!(m-\rho)^2 \rho^k}{k!} + (m-\rho)\rho^m \right]}. \end{aligned}$$

Если принять N_q в виде (4.24), то

$$T_q = \frac{p_0 \rho^{m+1}}{\lambda (m-1)!(m-\rho)^2} = \frac{p_0 \rho^m}{\mu (m-1)!(m-\rho)^2}. \quad (4.26)$$

4.3.4. Среднее время пребывания требования в системе М/М/м

Среднее время T_s пребывания одного требования в системе М/М/м складывается из среднего времени T_q пребывания в очереди и среднего времени обслуживания. Среднее время обслуживания обратно пропорционально интенсивности обслуживания μ , т. е. $1/\mu$. Тогда

$$T_s = T_q + \frac{1}{\mu} = \frac{p_0 \rho^m}{\mu(m-1)!(m-\rho)^2} + \frac{1}{\mu}. \quad (4.27)$$

4.3.5. Среднее число требований в системе М/М/м

Среднее число N_s требований в системе может быть рассчитано по формуле Литтла с учетом (4.27). Для этого умножим левую и правую части равенства (4.27) на λ , получим:

$$N_s = \lambda T_s = \frac{\lambda p_0 \rho^m}{\mu(m-1)!(m-\rho)^2} + \frac{\lambda}{\mu} = \frac{p_0 \rho^{m+1}}{(m-1)!(m-\rho)^2} + \rho. \quad (4.28)$$

Среднее число требований в системе можно выразить через среднее число требований в очереди:

$$N_s = N_q + \rho. \quad (4.29)$$

4.3.6. Среднее число свободных каналов обслуживания

Среднее число $N_{св}$ свободных каналов (приборов) обслуживания в системе М/М/м найдем как математическое ожидание случайной дискретной величины:

$$N_{св} = \sum_{k=0}^m k p_{m-k} = \sum_{k=1}^m k \frac{p_0 \rho^{(m-k)}}{(m-k)!} = p_0 \sum_{k=1}^m \frac{k \rho^{(m-k)}}{(m-k)!},$$

где m — число каналов обслуживания, p_0 определяется по формуле (4.21).

4.3.7. Среднее число занятых каналов обслуживания

Среднее число $N_{зан}$ занятых каналов обслуживания в системе М/М/м определим как математическое ожидание числа занятых каналов с учетом того, что в состояниях E_{m+i} , $i = 0, 1, \dots, \infty$ все каналы заняты:

$$N_{зан} = \sum_{k=0}^{m-1} k p_k + \sum_{i=0}^{\infty} m p_{m+i} = \sum_{k=0}^{m-1} k p_k + m \sum_{i=0}^{\infty} p_{m+i},$$

где вторая сумма представляет собой вероятность $P_{ож}$, которая может быть рассчитана по формулам (4.22), (4.23). Тогда

$$N_{зан} = \sum_{k=0}^{m-1} k p_k + m P_{ож} = \sum_{k=0}^{m-1} k \frac{p_0 \rho^k}{k!} + m P_{ож}. \quad (4.30)$$

4.3.8. Функция распределения времени ожидания начала обслуживания для требования, прибывшего в систему М/М/м

Рассмотрим сначала поток требований, выходящий из системы при условии, что все каналы обслуживания постоянно заняты — предельный случай. Обслуживание каждого требования осуществляется по экспоненциальному закону с параметром μ . Следовательно, интервалы времени между выходящими из системы требованиями распределены экспоненциально. Общий или результирующий поток требований, который формируется системой, состоящей из m обслуживающих каналов (приборов), складывается из суммы требований, которые покидают каждый из m канал обслуживания. Тогда можно считать, что параметр выходного результирующего потока будет равен $m\mu$, где m —

число каналов обслуживания. При этом выходной поток сформирован из стационарных потоков, поэтому он сам будет обладать свойством стационарности. Формирование выходного потока осуществляется через сумму независимых потоков, каждый из которых является ординарным и обладает отсутствием последействия. Рассмотрим, будет ли выполняться условие ординарности для выходного результирующего потока с параметром $m\mu$, т. е.

$$\lim_{t \rightarrow 0} \frac{P_{\geq 2}(t)}{t} = 0. \quad (4.31)$$

Для нахождения предельного отношения (4.31) определим вероятность $P_{\geq 2}(t)$.

Вероятность $P_{\geq 2}(t)$ поступления не менее двух требований на отрезке времени t из выходного пуассоновского потока определится в виде

$$\begin{aligned} P_{\geq 2}(t) &= \sum_{k=2}^{\infty} P_k(t) = \sum_{k=2}^{\infty} \frac{(m\mu t)^k}{k!} e^{-m\mu t} = e^{-m\mu t} \sum_{k=2}^{\infty} \frac{(m\mu t)^k}{k!} = \\ &= e^{-m\mu t} \left[\sum_{k=2}^{\infty} \frac{(m\mu t)^k}{k!} + \frac{(m\mu t)^0}{0!} + \frac{(m\mu t)}{1!} - \frac{(m\mu t)^0}{0!} - \frac{(m\mu t)}{1!} \right] = \\ &= e^{-m\mu t} \left[\sum_{k=0}^{\infty} \frac{(m\mu t)^k}{k!} - 1 - m\mu t \right] = e^{-m\mu t} [e^{m\mu t} - 1 - m\mu t] = \\ &= e^{-m\mu t} e^{m\mu t} - e^{-m\mu t} - m\mu t e^{-m\mu t} = 1 - e^{-m\mu t} - m\mu t e^{-m\mu t}. \end{aligned}$$

Предельное отношение $P_{\geq 2}(t)$ к t при $t \rightarrow 0$ имеет вид

$$\lim_{t \rightarrow 0} \frac{1 - e^{-m\mu t} - m\mu t e^{-m\mu t}}{t}.$$

Вычислим предел, используя правило Лопиталья (беря производные по t от числителя и знаменателя дроби и затем переходя к пределу):

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{m\mu e^{-m\mu t} - [m\mu e^{-m\mu t} - (m\mu)^2 t e^{-m\mu t}]}{1} &= \lim_{t \rightarrow 0} \{m\mu e^{-m\mu t} [1 - (1 - m\mu t)]\} = \\ &= \lim_{t \rightarrow 0} \{m\mu e^{-m\mu t} [1 - 1 + m\mu t]\} = \lim_{t \rightarrow 0} \{(m\mu)^2 t e^{-m\mu t}\} = 0. \end{aligned}$$

Как видим, предел отношения равняется нулю. Это значит, что выходной результирующий поток требований является ординарным.

Проверим выходной поток на отсутствие последействия. Напомним, что поток требований называется потоком без последействия, если условные вероятности поступления k требований на произвольном отрезке времени $[t_0; t_0+t]$, вычисленные при различных предположениях о распределении моментов поступления требований до t_0 , совпадают с безусловной вероятностью $P_k(t_0, t)$ [11]. В случае стационарного потока начало отсчета времени t_0 не имеет значения и его можно положить равным нулю. Для стационарного потока без последействия из формулы полной вероятности следует условие:

$$P_k(t + \Delta t) = \sum_{j=0}^k P_j(t) P_{k-j}(\Delta t), \quad (4.32)$$

где вероятности, входящие в (4.32), относятся к выходному потоку и поэтому имеют распределение Пуассона с параметром $m\mu t$ [11, с.12].

Проверим выполнение равенства (4.32). Для этого рассмотрим, левые и правые части равенства (4.32):

$$\begin{aligned} P_k(t + \Delta t) &= \frac{[m\mu(t + \Delta t)]^k}{k!} e^{-m\mu(t + \Delta t)}, \\ \sum_{j=0}^k P_j(t) P_{k-j}(\Delta t) &= \sum_{j=0}^k \frac{(m\mu t)^j e^{-m\mu t}}{j!} \cdot \frac{(m\mu \Delta t)^{k-j} e^{-m\mu \Delta t}}{(k-j)!} = \\ &= e^{-m\mu(t + \Delta t)} \sum_{j=0}^k \frac{(m\mu t)^j (m\mu \Delta t)^{k-j}}{j! (k-j)!}. \end{aligned}$$

Последнее соотношение умножим и разделим на $k!$, тогда получим:

$$\begin{aligned} \sum_{j=0}^k P_j(t) P_{k-j}(\Delta t) &= e^{-m\mu(t+\Delta t)} \sum_{j=0}^k \frac{k!(m\mu t)^j (m\mu \Delta t)^{k-j}}{k! j! (k-j)!} = \\ &= \frac{e^{-m\mu(t+\Delta t)}}{k!} \sum_{j=0}^k \frac{k!(m\mu t)^j (m\mu \Delta t)^{k-j}}{j! (k-j)!}. \end{aligned}$$

Сумма в последнем выражении представляет собой бином Ньютона. Напомним его для слагаемых a и b :

$$(a+b)^k = \sum_{j=0}^k \frac{k! a^j b^{k-j}}{j! (k-j)!}.$$

Следовательно, будем иметь:

$$\sum_{j=0}^k P_j(t) P_{k-j}(\Delta t) = \frac{e^{-m\mu(t+\Delta t)} (m\mu t + m\mu \Delta t)^k}{k!} = \frac{e^{-m\mu(t+\Delta t)} [m\mu(t+\Delta t)]^k}{k!}.$$

Таким образом, получили, что левая часть равенства (4.32) равна правой части. Тем самым доказали, что выходной поток без последствия. При этом параметр потока равен $m\mu$.

Обратимся к выводу функции распределения времени ожидания начала обслуживания. Рассмотрим вероятность $P_{m+k}(>t)$ того, что требование, прибывшее в систему, когда в ней уже находятся $m+k$ требований, прождет в очереди больше времени, чем время t . Для этого необходимо, чтобы за время t систему покинуло не более k требований. Вероятность $\bar{P}_{\leq m+k}(t)$ того, что систему за время t покинет не более k требований найдем по формуле сложения вероятностей для пуассоновского потока:

$$\bar{P}_{\leq m+k}(t) = \sum_{j=0}^k \frac{(m\mu t)^j}{j!} e^{-m\mu t} \quad (4.33)$$

В формуле (4.33) вероятность $\bar{P}_{\leq m+k}(t)$ можно трактовать как вероятность того, что на отрезке времени t в пуассоновском потоке с параметром $m\mu$ будет не более $m+k$ требований. Вероятности $P_{m+k}(>t)$ и $\bar{P}_{\leq m+k}(t)$ равны

между собой, поскольку описывают равнозначные события. Теперь рассмотрим вероятность $P(> t)$ того, что всякое требование, поступившее в систему М/М/т, прождет в очереди больше времени, чем время t . Эту вероятность следует рассчитывать с учетом вероятности имеющихся требований в системе, т. е. по формуле полной вероятности:

$$P(> t) = \sum_{k=0}^{\infty} p_{m+k} P_{m+k}(> t),$$

где p_{m+k} — стационарная вероятность того, что в системе находится $m+k$ требований. Вероятность p_{m+k} определяется выражением (4.19) с учетом значения нижнего индекса:

$$p_{m+k} = \frac{q^{m+k} m^m p_0}{m!}.$$

Тогда, учитывая, что $P_{m+k}(> t) = \bar{P}_{\leq m+k}(t)$, имеем

$$\begin{aligned} P(> t) &= \sum_{k=0}^{\infty} p_{m+k} P_{m+k}(> t) = \sum_{k=0}^{\infty} \frac{q^{m+k} m^m p_0}{m!} \sum_{j=0}^k \frac{(m\mu t)^j}{j!} = \\ &= \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} q^k \sum_{j=0}^k \frac{(m\mu t)^j}{j!}, \end{aligned}$$

где $q = \frac{\lambda}{m\mu} < 1$.

Выполним преобразования для последних двух сумм:

$$\begin{aligned} P(> t) &= \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} q^k \sum_{j=0}^k \frac{(m\mu t)^j}{j!} = \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{(m\mu t)^j}{j!} q^k q^j q^{-j} = \\ &= \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{(m\mu q t)^j}{j!} q^{(k-j)}. \end{aligned}$$

Теперь можно применить правило Коши, которое в классическом случае имеет следующий вид [17, с. 133]:

$$\left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{k=0}^{\infty} b_k \right) = \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k}.$$

Итак, в соответствии с правилом Коши будем иметь

$$P(> t) = \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} \frac{(m\mu q t)^k}{k!} \sum_{k=0}^{\infty} q^k. \quad (4.34)$$

В (4.34) последняя сумма — это бесконечно убывающая геометрическая прогрессия, поскольку $q < 1$. Преобразуем (4.34) с учетом формулы бесконечной суммы убывающей геометрической прогрессии:

$$P(> t) = \frac{q^m m^m p_0 e^{-m\mu t}}{m!} \sum_{k=0}^{\infty} \frac{(m\mu t)^k q^k}{k!} \cdot \frac{1}{1-q}$$

или

$$\begin{aligned} P(> t) &= \frac{q^m m^m p_0 e^{-m\mu t}}{m!(1-q)} \sum_{k=0}^{\infty} \frac{(m\mu t)^k (\lambda/m\mu)^k}{k!} = \frac{q^m m^m p_0 e^{-m\mu t}}{m!(1-q)} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = \\ &= \frac{q^m m^m p_0 e^{-m\mu t} e^{\lambda t}}{m!(1-q)}, \end{aligned}$$

где бесконечная сумма была заменена экспонентой с показателем λt .

Окончательно получаем

$$\begin{aligned} P(> t) &= \frac{q^m m^m p_0 e^{-m\mu t} e^{\lambda t}}{m!(1-q)} = p_0 \frac{(\lambda m / m\mu)^m e^{-(m\mu-\lambda)t}}{m!(1-\lambda/m\mu)} = p_0 \frac{\rho^m m e^{-(m\mu-\lambda)t}}{m!(m-\rho)} = \\ &= p_0 \frac{\rho^m e^{-(m\mu-\lambda)t}}{(m-1)!(m-\rho)}, \end{aligned} \quad (4.35)$$

где $\rho = \lambda / \mu$.

Выражение (4.35) определяет вероятность того, что требование, прибывшее в систему, будет ждать начала обслуживания больше чем время t . Вероятность $P(T < t)$ того, что требование, прибывшее в систему, будет ждать начала обслуживания не меньше, чем время t , можно найти из полной группы событий:

$$P(> t) + P(T \leq t) = 1, \Rightarrow P(T \leq t) = 1 - P(> t).$$

Поскольку вероятность $P(T \leq t)$ — это есть функция распределения времени ожидания начала обслуживания $F_q(t)$, то при использовании противоположной вероятности $P(> t)$, получим

$$F_q(t) = 1 - P(> t) = 1 - p_0 \frac{\rho^m e^{-(m\mu - \lambda)t}}{(m-1)!(m-\rho)}. \quad (4.36)$$

Характерный вид функции распределения времени ожидания начала обслуживания показан на рис. 4.4.

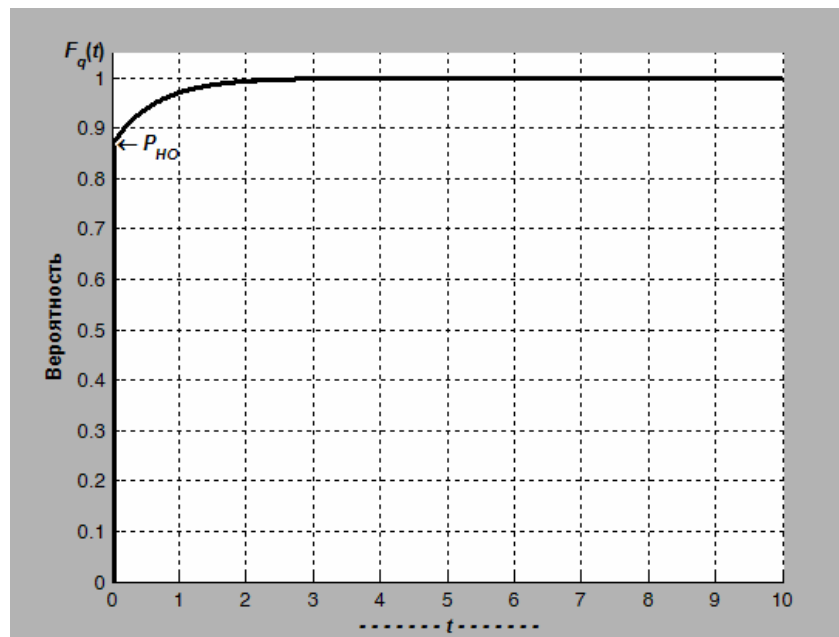


Рис. 4.4. Функция распределения времени ожидания начала обслуживания в системе M/M/7 с параметрами $\lambda = 2, \mu = 0.5, m = 7$

Характер поведения функции $F_q(t)$ на рис. 4.4 объясняется следующим образом. Начальный скачок функции $F_q(t)$ соответствует вероятности того, что поступившее в систему требование не будет ждать начала обслуживания. Ранее была определена вероятность ожидания начала обслуживания, которая вычисляется по формуле (4.22), т. е.

$$P_{о.ж} = p_0 \frac{(mq)^m}{m!(1-q)}, \text{ где } p_0 = \left[\sum_{k=0}^{m-1} \frac{(mq)^k}{k!} + \frac{(mq)^m}{m!(1-q)} \right]^{-1}, \quad q = \frac{\lambda}{m\mu}.$$

Для системы М/М/7 с параметрами $\lambda = 2$, $\mu = 0.5$, $m = 7$ получим следующие значения $P_{ож}$ и p_0 : $P_{ож} = 0.1351$, $p_0 = 0.0178$.

Тогда из условия полной группы событий можно вычислить вероятность P_{HO} того, что ждать не придется:

$$P_{ож} + P_{HO} = 1, \Rightarrow P_{HO} = 1 - P_{ож} = 1 - 0.1351 = 0.8649.$$

Кроме того, если в формулу (4.36) ввести $q = \lambda / m\mu$, то получим следующее:

$$F_q(t) = 1 - \frac{p_0 \rho^m e^{-(m\mu - \lambda)t}}{(m-1)!(m-\rho)} = 1 - \frac{p_0 (m\lambda / m\mu)^m e^{-(m\mu - \lambda)t}}{(m-1)!m(1-\rho/m)} = 1 - \frac{p_0 (mq)^m e^{-(m\mu - \lambda)t}}{m!(1-q)},$$

откуда видно, что функцию распределения $F_q(t)$ можно представить в виде

$$F_q(t) = 1 - P_{ож} e^{-(m\mu - \lambda)t}. \quad (4.37)$$

4.4. Система М/М/м с отказами. В-формула Эрланга

Условие работы системы таково: если требование поступает в момент времени, когда все m приборов (каналов, линий, узлов и т.п.) заняты, то оно теряется. В системе М/М/м с отказами очередь отсутствует и поэтому система имеет конечное множество состояний, число которых определяется числом приборов обслуживания m . Описание функционирования системы, как и ранее, сделаем в терминах процесса размножения и гибели при следующих параметрах:

$$\lambda_k = \begin{cases} \lambda = \text{const}, & k < m; \\ 0, & k \geq m; \end{cases} \quad \begin{aligned} \mu_k &= k\mu, & k &= 1, 2, \dots, m, \\ \mu &= \text{const}. \end{aligned} \quad (4.38)$$

Размеченный граф состояний системы М/М/м с параметрами (4.38) показан на рис. 4.5.

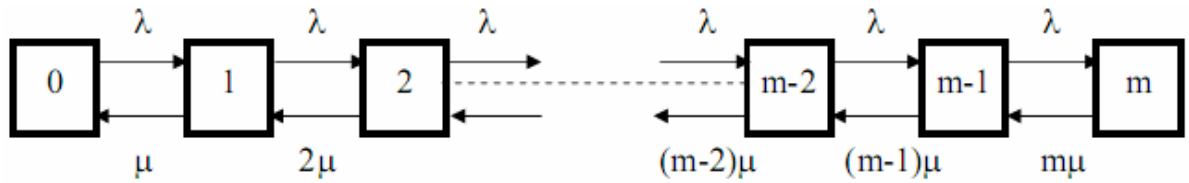


Рис. 4.5. Граф состояний системы М/М/м с отказами

Назначение параметров системы, показанной на рис. 4.5, может быть объяснено следующим образом. Во-первых, поток требований последовательно переводит систему из любого данного состояния в правое соседнее состояние с одной и той же интенсивностью λ . Перевод системы из какого-либо данного состояния в левое соседнее состояние осуществляется за счет обслуживания требований и можно говорить об интенсивности потока обслуживаний, который уже зависит от исходного данного состояния. Например, если система находится в состоянии 2 (см. рис. 4.5), когда два канала заняты, то она может перейти в состояние 1 в том случае, если закончит обслуживание либо первый, либо второй канал, поэтому суммарная интенсивность их потоков обслуживания будет равна 2μ . Интенсивность обслуживания увеличивается, поскольку может освободиться любой из двух каналов обслуживания. Таким образом, суммарный поток обслуживаний будет иметь интенсивность, кратную исходному данному состоянию.

Вероятности состояний цепи, приведенной на рис. 4.5, рассчитаем по общей формуле (2.46) с учетом заданных параметров (4.38):

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} = p_0 \left[\frac{\lambda}{\mu} \cdot \frac{\lambda}{2\mu} \cdot \frac{\lambda}{3\mu} \cdot \dots \cdot \frac{\lambda}{k\mu} \right] = p_0 \left(\frac{\lambda}{\mu} \right)^k \cdot \frac{1}{k!}$$

или, полагая $\rho = \lambda / \mu$, получим

$$p_k = \begin{cases} p_0 \frac{\rho^k}{k!}, & k \leq m; \\ 0, & k > m. \end{cases} \quad (4.39)$$

Вероятность p_0 определим по формуле (2.47) с учетом параметров (4.38):

$$p_0 = \frac{1}{1 + \sum_{k=1}^m \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu}} = \frac{1}{1 + \sum_{k=1}^m \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k} = \frac{1}{1 + \sum_{k=1}^m \frac{\rho^k}{k!}} = \frac{1}{\sum_{k=0}^m \frac{\rho^k}{k!}}. \quad (4.40)$$

Для случая $k = m$ по формуле (4.39) с учетом (4.40) получаем широко применяемую в телефонии так называемую *B*-формулу Эрланга [4]:

$$p_m = \frac{\rho^m / m!}{\sum_{k=0}^m \rho^k / k!} = \frac{(\lambda / \mu)^m / m!}{\sum_{k=0}^m (\lambda / \mu)^k / k!}. \quad (4.41)$$

Вероятность p_m описывает долю времени, когда все m приборов заняты. В телефонии по формуле (4.41) созданы таблицы и номограммы, которые используются для анализа работоспособности телефонных сетей. Часто формулы (4.39), (4.40) также называют формулами Эрланга по имени датского инженера А. К. Эрланга, который в 20-х гг. XX в. впервые исследовал систему обслуживания с отказами применительно к телефонной связи [18]. Задачу анализа многоканальной системы с отказами в обслуживании называют еще *классической задачей Эрланга*.

Математическая модель системы М/М/м с отказами в виде дифференциальных уравнений Колмогорова для процесса размножения и гибели с параметрами (4.38) будет иметь следующий вид:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda P_0(t) + \mu P_1(t), \\ \frac{dP_1(t)}{dt} &= -(\lambda + \mu)P_1(t) + \lambda P_0(t) + 2\mu P_2(t), \\ \frac{dP_2(t)}{dt} &= -(\lambda + 2\mu)P_2(t) + \lambda P_1(t) + 3\mu P_3(t), \\ &\dots\dots\dots \\ \frac{dP_{m-1}(t)}{dt} &= -[\lambda + (m-1)\mu]P_{m-1}(t) + \lambda P_{m-2}(t) + m\mu P_m(t), \\ \frac{dP_m(t)}{dt} &= \lambda P_{m-1}(t) - m\mu P_m(t). \end{aligned} \quad (4.42)$$

Систему дифференциальных уравнений (4.42) легко можно получить по заданному графу состояний рис. 4.5 в соответствии с мнемоническим правилом: производная вероятности данного состояния равна алгебраической сумме потоков вероятностей, проходящих через воображаемую окружность вокруг данного состояния. Для решения системы (4.42) задаются естественные начальные условия:

$$P_0(0) = 1, \quad P_i(0) = 0, \quad i = 1, 2, \dots, m. \quad (4.43)$$

Вероятности $P_k(t)$, $k = \overline{1, m}$, характеризуют среднюю загрузку системы, и ее изменение с течением времени [9].

4.5. Операционные характеристики системы М/М/т с отказами

Рассмотрим следующие характеристики системы М/М/т с отказами: вероятность отказа в обслуживании $p_{отк}$, относительная пропускная способность Q , абсолютная пропускная способность A , среднее число занятых обслуживанием каналов N_z , среднее число свободных каналов обслуживания $N_{св}$, коэффициент занятости каналов обслуживанием K_z , среднее время пребывания требования в системе T_c , среднее время простоя каналов обслуживания $T_{пр}$.

Определим $p_{отк}$. Требование, поступающее в систему, получает отказ в том случае, когда все каналы обслуживания заняты. Поэтому вероятность отказа в обслуживании $p_{отк}$ будет равна вероятности того, что заняты все m каналов. С учетом формулы (4.39) имеем

$$p_{отк} = p_m = p_0 \frac{\rho^m}{m!} = \frac{\rho^m / m!}{\sum_{k=0}^m \frac{\rho^k}{k!}}, \quad (4.44)$$

где $\rho = \lambda / \mu$, а p_0 определяется по формуле (4.40). Вероятность $p_{отк}$ характеризует долю требований, получающих отказ.

Относительная пропускная способность Q определяет долю обслуженных требований и соответственно Q — это вероятность того, что требование будет обслужено. Величину Q определим из полной группы событий. В системах с отказами требования или обслуживаются или получают отказ (очереди нет). Поэтому

$$Q + p_{отк} = 1, \Rightarrow Q = 1 - p_{отк} = 1 - p_0 \frac{\rho^m}{m!}. \quad (4.45)$$

Абсолютная пропускная способность A определяется как среднее число требований, обслуживаемых в единицу времени. В системе обслуживаются те требования, которые поступают в нее с интенсивностью λ . Средняя доля пришедших требований и обслуживаемых системой определяется как относительная пропускная способность Q системы. Поэтому величину абсолютной пропускной способности системы определяется в виде

$$A = \lambda Q = \lambda \left(1 - p_0 \frac{\rho^m}{m!} \right). \quad (4.46)$$

Если за единицу времени принять одну секунду (c), то единицей измерения абсолютной пропускной способности A является: $A = [\text{требования} / c]$.

Среднее число занятых обслуживанием каналов N_z определим как математическое ожидание дискретной случайной величины:

$$\begin{aligned} N_z &= \sum_{k=0}^m k p_k = \sum_{k=1}^m k p_k = \sum_{k=1}^m \frac{k p_0 \rho^k}{k!} = \sum_{k=1}^m \frac{p_0 \rho^k}{(k-1)!} = \rho \sum_{k=1}^m \frac{p_0 \rho^{k-1}}{(k-1)!} = \\ &= \left\{ k-1 = j, \quad k=1, j=0, \quad k=m, j=m-1 \right\} = \rho \sum_{j=0}^{m-1} \frac{p_0 \rho^j}{j!}. \end{aligned}$$

Сумма в последнем выражении определяет собой вероятность того, что в системе будет не более $(m-1)$ требований. Если к этой сумме добавить вероятность нахождения в системе точно m требований, то получим полную группу событий и, следовательно:

$$\sum_{j=0}^{m-1} \frac{p_0 \rho^j}{j!} + \frac{p_0 \rho^m}{m!} = 1, \Rightarrow \sum_{j=0}^{m-1} \frac{p_0 \rho^j}{j!} = 1 - \frac{p_0 \rho^m}{m!}. \quad (4.47)$$

С учетом (4.47) получим выражение для N_3 :

$$N_3 = \rho \left(1 - \frac{p_0 \rho^m}{m!} \right). \quad (4.48)$$

Среднее число свободных каналов обслуживания $N_{св}$ определим как разность между общим числом каналов системы и средним числом занятых обслуживанием каналов:

$$N_{св} = m - N_3 = m - \rho \left(1 - \frac{p_0 \rho^m}{m!} \right). \quad (4.49)$$

Коэффициент занятости каналов обслуживанием K_3 определяется отношением среднего числа занятых каналов к их общему числу:

$$K_3 = \frac{N_3}{m} = \frac{\rho}{m} \left(1 - \frac{p_0 \rho^m}{m!} \right). \quad (4.50)$$

Среднее время T_c пребывания требования в системе М/М/м с отказами определим по формуле Литтла:

$$N_3 = \lambda T_c, \Rightarrow T_c = \frac{N_3}{\lambda} = \frac{\rho}{\lambda} \left(1 - \frac{p_0 \rho^m}{m!} \right) = \frac{1}{\mu} \left(1 - \frac{p_0 \rho^m}{m!} \right). \quad (4.51)$$

Среднее время простоя каналов обслуживания T_{np} также определим по формуле Литтла:

$$N_{св} = \lambda T_{np}, \Rightarrow T_{np} = \frac{N_{св}}{\lambda} = \frac{m}{\lambda} - \frac{\rho}{\lambda} \left(1 - \frac{p_0 \rho^m}{m!} \right) = \frac{m}{\lambda} - \frac{1}{\mu} \left(1 - \frac{p_0 \rho^m}{m!} \right), \quad (4.52)$$

где

$$\rho = \frac{\lambda}{\mu}, \quad p_0 = \frac{1}{\sum_{k=0}^m \frac{\rho^k}{k!}}.$$

4.6. Пример анализа автоматической телефонной станции как системы массового обслуживания с отказами

Условие примера. Автоматическая телефонная станция (АТС) обеспечивает не более 97 переговоров одновременно. Средняя продолжительность разговора равна 1 мин, а вызовы поступают в среднем через каждые 0.5 с. Рассматривая такую АТС как многоканальную систему обслуживания с отказами и простейшим входным потоком, определить:

- вероятность отказа $p_{отк}$;
- относительную пропускную способность Q ;
- абсолютную пропускную способность A ;
- среднее число занятых каналов N_z ;
- среднее время T_{cp} пребывания вызова на станции с учетом того, что разговор может и не состояться.

Характеристики рассматриваемой СМО:

- число каналов обслуживания $m = 97$;
- интенсивность входного потока $\lambda = 1/0.5 = 2 \text{ с}^{-1}$;
- интенсивность обслуживания $\mu = 1/60 = 0.01667 \text{ с}^{-1}$;
- приведенная плотность потока заявок $\rho = \lambda/\mu = 2/(1/60) = 120$.

Вероятности состояний системы с отказами в стационарном режиме:

$$p_k = p_0 \frac{\rho^k}{k!}, \quad k = \overline{1, 97}; \quad p_0 = \left(\sum_{k=0}^{97} \frac{\rho^k}{k!} \right)^{-1},$$

где p_k — вероятность того, что в системе будет точно k вызовов, p_0 — вероятность того, что в системе не будет вызовов.

Принимаемые допущения:

1) При больших значениях m ($m > 20$) полагают:

$$p_k = \frac{\rho^k}{k!} e^{-\rho}, \quad k = \overline{1, m}.$$

Это вытекает из приближенного разложения экспоненты, при замене бесконечного ряда на конечный:

$$p_0 = \left(\sum_{k=0}^{97} \frac{\rho^k}{k!} \right)^{-1} = \left(1 + \rho + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \dots + \frac{\rho^{97}}{97!} \right)^{-1} = e^{-\rho}.$$

2) При больших значениях приведенной плотности потока ρ (больше 20) применяется формула:

$$\sum_{k=0}^i \frac{\rho^k}{k!} e^{-\rho} \approx 0.5 + \Phi \left(\frac{i + 0.5 - \rho}{\sqrt{\rho}} \right),$$

где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt \text{ — функция Лапласа или интеграл вероятности.}$$

Свойство функции Лапласа: $\Phi(-x) = -\Phi(x)$. Значения $\Phi(x)$ табулированы и приводятся в учебниках и справочниках по теории вероятностей.

Расчет функции Лапласа в системе MATLAB:

Функция Лапласа может быть найдена как решение следующего дифференциального уравнения:

$$\frac{d\Phi}{dx} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \Phi(0) = 0.$$

Для решения дифференциального уравнения относительно функции Лапласа может быть использована следующая программа в системе MATLAB.

Создаем *M*-файл под именем Lap в следующем виде:

```
function Lap;  
clc  
L = 2;  
M = 1/60;  
p = L/M;  
m = 97;  
x = (m+0.5-p)/sqrt(p);  
if x<0  
    x = abs(x);  
F0 = 0;  
x0 = [0, x];  
[x1,F1] = ode45(@fun,x0,F0);  
R1 = 0.5 - F1(end)  
else  
F0 = 0;  
x0 = [0, x];  
[x1,F1] = ode45(@fun,x0,F0);  
R1 = 0.5 + F1(end)  
end  
  
%% M-функция описания правой части дифференциального  
уравнения
```

```
function f = fun(x, F)  
f = (1/(sqrt(2*pi))) * exp(-x^2/2);
```

Результат выполнения программы:

R1 =
0.0200

В рассматриваемом примере значение аргумента функции Лапласа отрицательное.

Величина R1 может быть вычислена также по следующему выражению:

$$R1 = \sum_{k=0}^{97} \frac{\rho^k}{k!} e^{-\rho} \approx 0.5 + \Phi\left(\frac{97 + 0.5 - 120}{\sqrt{120}}\right) = 0.5 + \Phi(-2.054) = 0.5 - \Phi(2.54) = 0.5 - 0.48 = 0.02.$$

Представление формул расчета вероятностей через специальные обозначения:

$$p_k = \frac{\rho^k}{k!} e^{-\rho} = \sum_{i=0}^k \frac{\rho^i}{i!} e^{-\rho} - \sum_{i=0}^{k-1} \frac{\rho^i}{i!} e^{-\rho} = \frac{\sum_{i=0}^k \frac{\rho^i}{i!} e^{-\rho} - \sum_{i=0}^{k-1} \frac{\rho^i}{i!} e^{-\rho}}{\sum_{i=0}^m \frac{\rho^i}{i!} e^{-\rho}},$$

где при больших m ($m > 20$) имеем

$$\sum_{i=0}^m \frac{\rho^i}{i!} e^{-\rho} = e^{\rho} e^{-\rho} = 1.$$

Применим следующие обозначения:

$$R(k, \rho) = \sum_{i=0}^k \frac{\rho^i}{i!} e^{-\rho} \approx 0.5 + \Phi \left(\frac{k + 0.5 - \rho}{\sqrt{\rho}} \right);$$

$$R(k-1, \rho) = \sum_{i=0}^{k-1} \frac{\rho^i}{i!} e^{-\rho} \approx 0.5 + \Phi \left(\frac{(k-1) + 0.5 - \rho}{\sqrt{\rho}} \right);$$

$$R(m, \rho) = \sum_{i=0}^m \frac{\rho^i}{i!} e^{-\rho} \approx 0.5 + \Phi \left(\frac{m + 0.5 - \rho}{\sqrt{\rho}} \right).$$

С учетом введенных обозначений расчет вероятностей состояний может производиться в виде

$$p_k = \frac{R(k, \rho) - R(k-1, \rho)}{R(m, \rho)}, \quad k = \overline{1, m}.$$

Расчет через функции Лапласа дает более точный результат по сравнению с заданным возведением в степень, вычислением факториалов, обращением экспоненциальной функции с большим показателем и т. д.

Расчет операционных характеристик АТС

1. Относительная пропускная способность Q :

$$Q = 1 - p_{отк} = 1 - p_m,$$

где вероятность отказа $p_{отк}$ равна вероятности занятости всех каналов p_m .

Рассчитаем вероятность p_m по формулам с использованием функции Лапласа:

$$p_m = \frac{R(m, \rho) - R(m-1, \rho)}{R(m, \rho)} = 1 - \frac{R(m-1, \rho)}{R(m, \rho)} = 1 - \frac{R(96, 120)}{R(97, 120)},$$

$$\begin{aligned} R(96, 120) &= 0.5 + \Phi \left(\frac{96 + 0.5 - 120}{\sqrt{120}} \right) = 0.5 + \Phi(-2.1452) = 0.5 - \Phi(2.1452) = \\ &= 0.5 - 0.484 = 0.016. \end{aligned}$$

Программа расчета $R(96,120)$ в MATLAB:

```
function Lap96  
L = 2;  
M = 1/60;  
p = L/M;  
m = 97;  
xx = (m-1+0.5-p)/sqrt(p);  
xx0 = [0, xx];  
  
if xx < 0  
    xx = abs(xx);  
[x2,F2] = ode45(@fun,xx0,F0);  
R2 = 0.5 + F2(end)  
end  
  
%% М-функция описания правой части дифференциального  
уравнения  
function f = fun(x, F)  
f = (1/(sqrt(2*pi)))*exp(-x^2/2);
```

Результат выполнения программы:

R2 =

0.0160

Значение $R(97,120)$ было рассчитано ранее как $R1 = 0.02$.

Таким образом, вероятность отказа и относительная пропускная способность равны

$$p_m = p_{отк} = 1 - \frac{0.016}{0.02} = 1 - 0.8 = 0.2;$$
$$Q = 1 - p_m = 1 - p_{отк} = 1 - 0.2 = 0.8.$$

2. Абсолютная пропускная способность A :

$$A = \lambda \cdot Q = 2 \cdot 0.8 = 1.6 \text{ вызовы/с.}$$

Величину A можно интерпретировать как среднее число вызовов, обслуженных в единицу времени.

3. Среднее число занятых каналов N_z

Величина N_{cp} может быть определена как отношение абсолютной производительности A системы (среднее число вызовов, обслуженных в единицу времени) к интенсивности обслуживания μ (среднее число вызовов, обслуживаемых в единицу времени одним каналом):

$$N_z = \frac{A}{\mu} = \frac{1.6}{1/60} = 96. \quad \left\{ N_{cp} = \frac{A}{\mu} = \frac{\lambda Q}{\mu} = \rho Q \right\}.$$

4. Среднее время T_{cp} пребывания вызова на станции с учетом того, что разговор может и не состояться

Значение T_{cp} вычислим по формуле Литтла:

$$N_z = \lambda T_{cp}; \quad \Rightarrow \quad T_{cp} = \frac{N_z}{\lambda} = \frac{96}{2} = 48 \text{ (с)}.$$

4.7. Система М/М/т/К с конечным накопителем

Рассматриваемая система имеет ограниченное число мест ожидания в очереди и m параллельно работающих приборов (каналов). Общее

число требований, находящихся на обслуживании и в очереди, составляет величину K . Входящий поток требований пуассоновский, обслуживание осуществляется по экспоненциальному закону. Если очередное требование приходит, когда система занята (заняты все каналы обслуживания плюс накопитель), то оно получает отказ в обслуживании и сразу покидает систему.

Для описания работы системы $M/M/m/K$ применим процесс размножения и гибели со следующими параметрами:

$$\begin{aligned}\lambda_k &= \lambda = \text{const}, \quad k = 0, 1, 2, \dots, K-1; \\ \mu_k &= k\mu, \quad k = 1, 2, \dots, m, \\ \mu_k &= m\mu, \quad m \leq k \leq K, \quad \mu = \text{const}.\end{aligned}\tag{4.53}$$

Размеченный граф состояний системы $M/M/m/K$ показан на рис. 4.6.

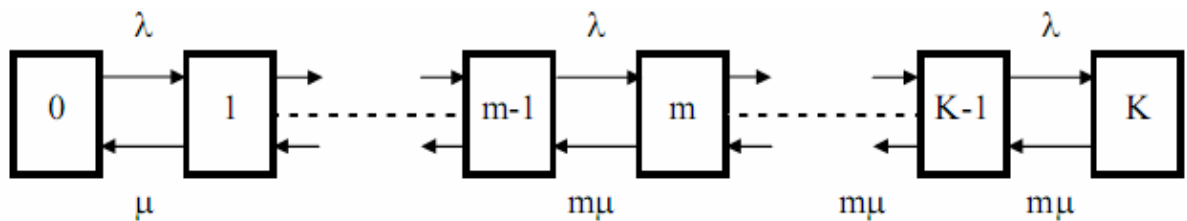


Рис. 4.6. Граф состояний системы $M/M/m/K$

В соответствии с графом состояний системы $M/M/m/K$ переход системы из любого состояния в правое соседнее состояние осуществляется за счет поступления требований с интенсивностью λ . Переход же системы из какого-либо состояния в левое соседнее состояние осуществляется за счет освобождения какого-либо канала обслуживания, причем интенсивность обслуживания пропорциональна числу работающих каналов: например, если система находится в состоянии E_3 (три требования находятся в трех каналах обслуживания), то переход в состояние E_2 (когда два требования находятся в двух каналах обслуживания, а значит и в системе) может произойти за счет освобождения любого из трех каналов (приборов) обслуживания, каждый из которых обладает интенсивностью обслуживания μ . Когда число

требований в системе равно или превышает число каналов, то интенсивность обслуживания остается постоянной и равной $m\mu$.

Определим стационарные вероятности системы М/М/м/К, заранее предполагая, что в системе М/М/м/К существует стационарный режим. Условие существования стационарного режима выясним позднее. Для случая, когда состояние системы не превышает значения m , определим:

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} = p_0 \frac{\lambda}{\mu} \cdot \frac{\lambda}{2\mu} \cdot \dots \cdot \frac{\lambda}{k\mu} = p_0 \frac{\rho^k}{k!}, \quad (4.54)$$

где $\rho = \lambda / \mu$, $k \leq m$.

Для дальнейшего расчета, когда возникает очередь в системе, применим мнемоническое правило: на границе раздела двух состояний потоки вероятностей равны между собой. Рассмотрим сначала границу между состоянием m и состоянием $m+1$:

$$\lambda p_m = m\mu p_{m+1}, \quad p_{m+1} = \frac{\lambda p_m}{m\mu} = \frac{\rho p_m}{m} = \frac{\rho p_0 \rho^m}{m m!} = \frac{p_0 \rho^{m+1}}{m m!},$$

где вероятность p_m определена по (4.54) при $k = m$.

Для границы состояний $m+1$ и $m+2$ будем иметь

$$\lambda p_{m+1} = m\mu p_{m+2}, \quad p_{m+2} = \frac{\lambda p_{m+1}}{m\mu} = \frac{\rho p_0 \rho^{m+1}}{m m m!} = \frac{p_0 \rho^{m+2}}{m^2 m!}.$$

Таким образом, продолжая аналогичные расчеты, можно получить

$$p_{m+k} = \frac{p_0 \rho^{m+k}}{m^k m!}, \quad 0 \leq k \leq (K - m), \quad (4.55)$$

где $m + k$ не превышает допустимой величины K .

С учетом (4.54), (4.55) стационарные вероятности системы М/М/м/К равны:

$$p_k = \begin{cases} \frac{p_0 \rho^k}{k!}, & k \leq m, \\ \frac{p_0 \rho^k}{m^{k-m} m!}, & m \leq k \leq K. \end{cases} \quad (4.56)$$

Вероятность p_0 найдем из нормировочного условия, используя (4.55):

$$\begin{aligned} \sum_{k=0}^K p_k = 1, \quad \sum_{k=0}^m p_k + \sum_{k=m+1}^K p_k = 1, \quad \sum_{k=0}^m \frac{p_0 \rho^k}{k!} + \sum_{k=1}^{K-m} \frac{p_0 \rho^{m+k}}{m^k m!} = 1, \\ p_0 \left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{k=1}^{K-m} \frac{\rho^k}{m^k} \right] = 1, \quad p_0 = \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{k=1}^{K-m} \left(\frac{\rho}{m} \right)^k \right]}. \end{aligned}$$

К последнему выражению для расчета p_0 можно применить формулу суммы конечной геометрической прогрессии, которая в общем случае имеет вид

$$s_n = \sum_{j=0}^n a_0 r^j = a_0 \frac{1 - r^{n+1}}{1 - r}, \quad (4.57)$$

где a_0 — первый член прогрессии, $r \neq 1$ — знаменатель прогрессии [17].

Формулу (4.57) можно применить к расчету вероятности p_0 в том случае, если потребовать, чтобы $\rho/m < \infty$. Будем считать, что это требование выполнено. Тогда

$$\begin{aligned} p_0 &= \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{k=1}^{K-m} \left(\frac{\rho}{m} \right)^k \right]} = \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m \rho^{K-1-m}}{m! m} \sum_{k=0}^{K-1-m} \left(\frac{\rho}{m} \right)^k \right]} = \\ &= \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^{m+1}}{m m!} \cdot \frac{1 - (\rho/m)^{K-m}}{1 - (\rho/m)} \right]} = \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^{m+1}}{m!(m - \rho)} \left(1 - \left(\frac{\rho}{m} \right)^{K-m} \right) \right]}, \end{aligned}$$

где $(K - m)$ — допустимая длина очереди в системе М/М/м/К. Если допустимую длину очереди обозначить через S , то вероятность p_0 принимает вид

$$p_0 = \frac{1}{\left[\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^{m+1}}{m!(m-\rho)} \left(1 - \left(\frac{\rho}{m} \right)^S \right) \right]}. \quad (4.58)$$

Поскольку система М/М/м/К имеет конечное число состояний, то интенсивность входного потока требований может быть достаточно большой. Очередное требование получит отказ в обслуживании, если система будет заполненной до величины K требований. Это означает, что в системе М/М/м/К всегда существует стационарный режим, для которого должно выполняться только условие $\rho/m < \infty$. Стационарные вероятности могут быть вычислены по формулам (4.56), (4.58).

4.8. Операционные характеристики системы М/М/м/К

Вероятность отказа $p_{отк}$ равна вероятности того, что система будет полностью занятой, т. е. когда в ней K требований:

$$p_{отк} = \frac{p_0 \rho^K}{m^{K-m} m!} = \frac{p_0 \rho^{m+S}}{m^S m!}, \quad (4.59)$$

где S — допустимая длина очереди, p_0 — вероятность отсутствия требований в системе, рассчитывается по формуле (4.58).

Относительная пропускная способность Q :

$$Q = 1 - p_{отк} = 1 - \frac{p_0 \rho^{m+S}}{m^S m!}. \quad (4.60)$$

Абсолютная пропускная способность A :

$$A = \lambda Q = \lambda \left(1 - \frac{p_0 \rho^{m+S}}{m^S m!} \right). \quad (4.61)$$

Абсолютная пропускная способность — это среднее число требований, обслуживаемых в единицу времени (единица измерения c^{-1}).

Вероятность наличия очереди p_q равна:

$$p_q = \sum_{k=m+1}^{m+S} p_k = \sum_{k=1}^S \frac{p_0 \rho^{m+k}}{m^k m!} = \frac{p_0 \rho^m}{m!} \sum_{k=1}^S \left(\frac{\rho}{m} \right)^k.$$

Применяя формулу суммы конечной геометрической прогрессии, получим

$$p_q = \frac{p_0 \rho^m}{m!} \cdot \frac{1 - (\rho/m)^S}{1 - (\rho/m)} = \frac{p_0 \rho^{m+1}}{m! (m - \rho)} \left[1 - \left(\frac{\rho}{m} \right)^S \right]. \quad (4.62)$$

Вероятность того, что все каналы обслуживания заняты. Эта вероятность $p_{зан}$ рассчитывается для случая, когда в очереди 0 требований, 1 требование, ..., S требований:

$$p_{зан} = \sum_{k=m}^{m+S} p_k = \sum_{k=0}^S \frac{p_0 \rho^{m+k}}{m^k m!} = \frac{p_0 \rho^m}{m!} \sum_{k=0}^S \left(\frac{\rho}{m} \right)^k.$$

К последнему выражению применима классическая формула суммы конечной геометрической прогрессии, поэтому будем иметь:

$$p_{зан} = \frac{p_0 \rho^m}{m!} \left(\frac{1 - (\rho/m)^{S+1}}{1 - (\rho/m)} \right) = \frac{p_0 \rho^m}{m!} \left(\frac{m - \rho (\rho/m)^S}{(m - \rho)} \right). \quad (4.63)$$

Вероятность того, что в системе будет не более k требований:

1) $0 \leq k \leq m$

$$p_1(\leq k) = \sum_{i=0}^k p_i = \sum_{i=0}^k \frac{\rho^i}{i!} p_0, \quad (4.64)$$

2) $m \leq k \leq K$

$$p_2(\leq k) = \sum_{i=m}^K p_i = \sum_{i=m}^K \frac{p_0 \rho^{m+i}}{m! m^i} = \frac{p_0 \rho^m}{m!} \sum_{i=m}^K \left(\frac{\rho}{m} \right)^i.$$

Применяя к последней сумме формулу суммы конечной геометрической прогрессии, получим

$$p_2(\leq k) = \frac{p_0 \rho^{m+1}}{m!m} \cdot \frac{1 - (\rho/m)^K}{1 - (\rho/m)} = \frac{p_0 \rho^{m+1}}{m! (m - \rho)} \left(1 - \left(\frac{\rho}{m} \right)^K \right). \quad (4.65)$$

Если требуется вычислить вероятность того, что в системе будет не менее k требований, то это можно сделать из полной группы событий. Например, найдем вероятность $p_1(\geq k)$ того, что в системе будет не менее k требований при $0 \leq k \leq m$:

$$p_1(\geq k) + p_1(\leq k) = 1, \Rightarrow p_1(\geq k) = 1 - p_1(\leq k) = 1 - \sum_{i=0}^k \frac{p_0 \rho^i}{i!}.$$

Среднее число требований N_{cp} в системе определяется как математическое ожидание случайной дискретной величины:

$$\begin{aligned} N_{cp} &= \sum_{k=0}^K k p_k = \sum_{k=0}^m k \frac{p_0 \rho^k}{k!} + \sum_{k=1}^{K-m} k \frac{p_0 \rho^{m+k}}{m! m^k} = p_0 \sum_{k=1}^m \frac{\rho^k}{(k-1)!} + \frac{p_0 \rho^m}{m!} \sum_{k=1}^{K-m} k \left(\frac{\rho}{m} \right)^k = \\ &= p_0 \rho \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{p_0 \rho^m}{m!} \sum_{k=1}^{K-m} k \left(\frac{\rho}{m} \right)^k. \end{aligned}$$

Последнюю сумму найдем с помощью искусственного дифференцирования с последующим применением формулы суммы конечной геометрической прогрессии:

$$\begin{aligned} \sum_{k=1}^{K-m} k \left(\frac{\rho}{m} \right)^k &= \left\{ \frac{\rho}{m} = q, \quad K - m = S \right\} = \sum_{k=1}^S k q^k = q \frac{d}{dq} \sum_{k=1}^S q^k = q \frac{d}{dq} \left[\frac{q - q^{S+1}}{1 - q} \right] = \\ &= q \frac{[1 - (S+1)q^S](1 - q) + (q - q^{S+1})}{(1 - q)^2} = q \frac{1 - (S+1)q^S - q + (S+1)q^{S+1} + q - q^{S+1}}{(1 - q)^2} = \\ &= \frac{q - (S+1)q^{S+1} + (S+1)q^{S+2} - q^{S+2}}{(1 - q)^2} = \frac{q - (S+1)q^{S+1} + Sq^{S+2}}{(1 - q)^2} = \left\{ q = \frac{\rho}{m} \right\} = \end{aligned}$$

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

Величина N_{cp} будет равна

$$N_{cp} = p_0 \left(\rho \sum_{k=0}^{m-1} \frac{\rho^k}{k!} + \frac{\rho^{m+1}}{(m-1)!} \cdot \frac{1 - (S+1)\left(\frac{\rho}{m}\right)^S + S\left(\frac{\rho}{m}\right)^{S+1}}{(m - \rho)^2} \right). \quad (4.66)$$

Найдем среднюю длину очереди N_q :

$$\begin{aligned} N_q &= \sum_{k=0}^{K-m} k p_{m+k} = \sum_{k=0}^{K-m} k \frac{p_0 \rho^{m+k}}{m! m^k} = \frac{p_0 \rho^m}{m!} \sum_{k=0}^{K-m} k \left(\frac{\rho}{m}\right)^k = \left\{ K - m = S, \quad \frac{\rho}{m} = q \right\} = \\ &= \frac{p_0 \rho^m}{m!} q \frac{d}{dq} \sum_{k=0}^S q^k = \frac{p_0 \rho^m}{m!} q \frac{d}{dq} \left[\frac{1 - q^{S+1}}{1 - q} \right] = \frac{p_0 \rho^m}{m!} q \frac{-(S+1)q^S(1-q) + (1-q^{S+1})}{(1-q)^2} = \\ &= \frac{p_0 \rho^m}{m!} q \frac{1 - q^{S+1} - Sq^S - q^S + Sq^{S+1} + q^{S+1}}{(1-q)^2} = \frac{p_0 \rho^m}{m!} \cdot \frac{q - (S+1)q^{S+1} + Sq^{S+2}}{(1-q)^2} = \\ &= \left\{ q = \frac{\rho}{m} \right\} = \frac{p_0 \rho^m m^2 [(m+1)\rho - (S+1)m\rho^{S+1} + S\rho^{S+2}]}{m! m^{S+2} (m - \rho)^2} = \\ &= \frac{p_0 \rho^{m+1} [m+1 - (S+1)m\rho^S + S\rho^{S+1}]}{m! m^S (m - \rho)^2}. \end{aligned} \quad (4.67)$$

4.9. Система М/М/т с ограниченным временем ожидания

4.9.1. Анализ системы М/М/т с ограниченным временем ожидания в стационарном режиме

Рассмотрим систему с m параллельно работающими приборами, с неограниченной очередью и ограниченным временем ожидания в очереди. Требования, находящиеся в очереди, могут покинуть систему до начала обслуживания. Уход требований из очереди происходит вследствие превышения какого-либо времени ожидания. Примем, что время ожидания начала обслуживания является случайным и подчиняется экспоненциальному закону, например, заданным своей функцией плотности:

$$h(t) = \nu e^{-\nu t} \quad (t > 0), \quad (4.68)$$

где ν — интенсивность ухода из очереди, т. е. величина, обратная среднему времени ожидания.

Величину ν можно интерпретировать, как плотность "потока уходов" требования, находящегося в очереди [9]. Требования, уходящие из системы, не дождавись обслуживания, часто называют "нетерпеливыми" требованиями.

Изменение состояний системы происходит в случае поступления требований или за счет обслуживания требований, когда нет очереди. При наличии очереди, т. е. когда состояние системы определяется величиной не менее $m+1$, изменение состояния может происходить также и за счет ухода "нетерпеливых" требований. При установлении экспоненциального закона ухода из очереди можно считать, что поток уходов "нетерпеливых" требований является простейшим. В случае протекания

простейших процессов в системе легко построить размеченный граф состояний такой системы. Этот граф состояний показан на рис. 4.7.

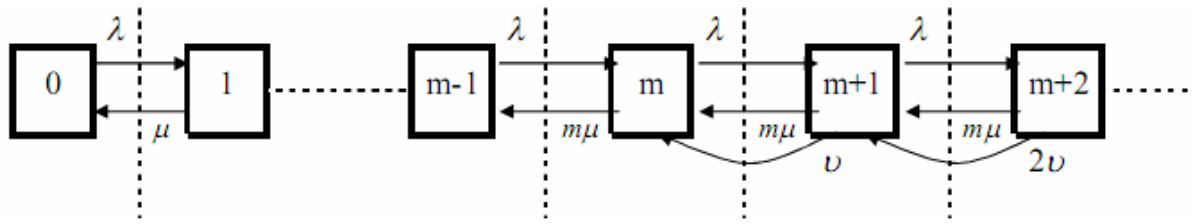


Рис. 4.7. Граф состояний системы М/М/м с ограниченным временем ожидания

На рис. 4.7 вертикальные штриховые линии между состояниями назовем границами между двумя соседними состояниями.

По размеченному графу состояний составим уравнения относительно стационарных вероятностей по известному мнемоническому правилу: потоки вероятностей между любыми двумя соседними состояниями равны [8, с.192]. Напомним, что поток вероятности — это произведение вероятности p_i (предельной или стационарной вероятности) состояния i , из которого исходит стрелка, на интенсивность потока событий, переводящего систему по этой стрелке. Теперь задача заключается в том, что бы подсчитать количество потоков вероятностей, проходящих через левую границу между двумя соседними состояниями и через правую границу, а затем уравнять их. Граф состояний системы М/М/м с ограниченным временем ожидания можно разделить на две части: состояния без очереди (от 0 до m) и состояния с очередью (от $m+1$ и более). Рассмотрим сначала состояния без очереди.

Составим уравнение потоков вероятностей между 0-м и 1-м состояниями:

$$p_0 \lambda = p_1 \mu, \Rightarrow p_1 = \frac{\lambda}{\mu} p_0.$$

Составим уравнение потоков вероятностей между 1-м и 2-м состояниями:

$$p_1 \lambda = p_2 2\mu, \Rightarrow p_2 = \frac{\lambda}{2\mu} p_1 = \frac{\lambda}{2\mu} \frac{\lambda}{\mu} p_0 = p_0 \frac{\lambda^2}{2\mu^2}.$$

Составим уравнение потоков вероятностей между 2-м и 3-м состояниями:

$$p_2 \lambda = p_3 3\mu, \Rightarrow p_3 = \frac{\lambda}{3\mu} p_2 = \frac{\lambda}{3\mu} \frac{\lambda^2}{2\mu^2} p_0 = p_0 \frac{\lambda^3}{2 \cdot 3\mu^3} = p_0 \frac{\lambda^3}{3!\mu^3}.$$

Как видно наблюдается закономерность в определении стационарных вероятностей для случаев, когда нет очереди в системе. Запишем общее выражение для расчета стационарной вероятности при $0 < k \leq m$:

$$p_k = p_0 \frac{\lambda^k}{k! \mu^k} = p_0 \frac{\rho^k}{k!}, \quad 0 < k \leq m, \quad (4.69)$$

где $\rho = \lambda / \mu$, p_0 — вероятность отсутствия в системе требований. В дальнейшем будем предполагать, что $\rho < m$.

По формуле (4.69) рассчитаем вероятность наличия в системе точно m требований:

$$p_m = p_0 \frac{\lambda^m}{m! \mu^m} = p_0 \frac{\rho^m}{m!}. \quad (4.70)$$

Перейдем к рассмотрению вероятностей состояний, когда в системе имеется очередь, когда состояние системы определяется значениями $m+1$, $m+2$,

Составим уравнение потоков вероятностей между m и $(m+1)$ состояниями:

$$\begin{aligned} p_m \lambda &= p_{m+1} m\mu + p_{m+1} \nu, \Rightarrow \\ p_{m+1} &= \frac{\lambda}{m\mu + \nu} p_m = \frac{\lambda}{m\mu + \nu} p_0 \frac{\lambda^m}{m! \mu^m} = p_0 \frac{\lambda^{m+1}}{m! \mu^m (m\mu + \nu)} = p_0 \frac{\lambda^{m+1}}{m! \mu^{m+1} (m + \nu / \mu)} = \\ &= p_0 \frac{\rho^{m+1}}{m! (m + \nu / \mu)}. \end{aligned}$$

Составим уравнение потоков вероятностей между $(m+1)$ -м и $(m+2)$ -м состояниями:

$$\begin{aligned} p_{m+1}\lambda &= p_{m+2}m\mu + p_{m+2}2\nu, \Rightarrow \\ p_{m+2} &= \frac{\lambda}{m\mu + 2\nu} p_{m+1} = \frac{\lambda}{m\mu + 2\nu} p_0 \frac{\lambda^{m+1}}{m! \mu^{m+1} (m + \nu/\mu)} = \\ &= p_0 \frac{\lambda^{m+2}}{m! \mu (m + 2\nu/\mu) \mu^{m+1} (m + \nu/\mu)} = p_0 \frac{\lambda^{m+2}}{m! \mu^{m+2} (m + 2\nu/\mu) (m + \nu/\mu)} = \\ &= p_0 \frac{\rho^{m+2}}{m! (m + \nu/\mu) (m + 2\nu/\mu)}. \end{aligned}$$

Составим уравнение потоков вероятностей между $(m+2)$ и $(m+3)$ состояниями:

$$\begin{aligned} p_{m+2}\lambda &= p_{m+3}m\mu + p_{m+3}3\nu, \Rightarrow \\ p_{m+3} &= \frac{\lambda}{m\mu + 3\nu} p_{m+2} = \frac{\lambda}{m\mu + 3\nu} p_0 \frac{\lambda^{m+2}}{m! \mu^{m+2} (m + \nu/\mu) (m + 2\nu/\mu)} = \\ &= p_0 \frac{\lambda^{m+3}}{m! \mu^{m+3} (m + \nu/\mu) (m + 2\nu/\mu) (m + 3\nu/\mu)} = \\ &= p_0 \frac{\rho^{m+3}}{m! (m + \nu/\mu) (m + 2\nu/\mu) (m + 3\nu/\mu)}. \end{aligned}$$

Проведенные расчеты показывают определенную закономерность определения стационарных вероятностей при наличии очереди и "нетерпеливых" требований. Запишем общее выражение для определения вероятности состояния при наличии очереди в системе с "нетерпеливыми" требованиями:

$$p_{m+s} = p_0 \frac{\rho^{m+s}}{m! \prod_{j=1}^s (m + j\nu/\mu)}, \quad s \geq 1. \quad (4.71)$$

В формулы (4.69), (4.71) входит пока неизвестная вероятность p_0 . Определим ее из нормировочного условия:

$$\sum_{k=0}^{\infty} p_k = 1, \quad \sum_{k=0}^m p_k + \sum_{k=m+1}^{\infty} p_k = 1, \quad (4.72)$$

где вероятности под знаком каждой из сумм определяются по (4.69) и (4.71).

Преобразуем нормировочное условие (4.72) с учетом (4.69), (4.71) для $k \leq m, s \geq 1$:

$$\begin{aligned} \sum_{k=0}^m p_0 \frac{\rho^k}{k!} + \sum_{s=1}^{\infty} p_0 \frac{\rho^{m+s}}{m! \prod_{j=1}^s (m + j\nu/\mu)} = 1, \Rightarrow \\ p_0 \left(\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{\rho^s}{\prod_{j=1}^s (m + j\nu/\mu)} \right) = 1, \Rightarrow \\ p_0 = \left(\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{\rho^s}{\prod_{j=1}^s (m + j\beta)} \right)^{-1}, \end{aligned} \quad (4.73)$$

где $\beta = \nu/\mu$ — среднее число уходов требования, стоящего в очереди, приходящиеся на среднее время обслуживания одного требования.

Расчет вероятностей по формулам (4.71), (4.73) затруднителен тем, что в них входят бесконечные суммы, которые не преобразуются подобно тому, как это было для геометрических прогрессий. Для грубой оценки ошибки, происходящей от отбрасывания членов сумм, начиная с r -го, можно пользоваться формулой [9]:

$$\sum_{s=r}^{\infty} \frac{\rho^s}{\prod_{j=1}^s (m + j\beta)} < \frac{\left(\frac{\rho}{\beta}\right)^r}{r!} e^{\left(\frac{\rho}{\beta}\right)}. \quad (4.74)$$

Важной характеристикой системы массового обслуживания является вероятность того, что требование (заявка) покинет систему не обслуженным P_n или вероятность отказа. Для системы М/М/м с ограниченным временем ожидания величину P_n определяют из следующих соображений: при установившемся режиме вероятность того, что заявка покинет систему не обслуженной, есть отношение среднего числа заявок, уходящих из очереди в единицу времени, к среднему числу заявок, поступающих в единицу времени [9, с.555]. Найдем среднее число заявок, уходящих из очереди в единицу времени. Для этого сначала вычислим математическое ожидание m_s числа заявок, находящихся в очереди, принимая во внимание формулы (4.71), (4.73):

$$m_s = M[s] = \sum_{s=1}^{\infty} s p_{m+s} = \frac{\frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{s \rho^s}{\prod_{j=1}^s (m + j\beta)}}{\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{\rho^s}{\prod_{j=1}^s (m + j\beta)}}. \quad (4.75)$$

Далее, чтобы получить вероятность P_n , нужно m_s умножить на среднюю "плотность уходов" одного требования ν и разделить на среднюю плотность требований λ , т. е. умножить на коэффициент

$$\frac{\nu}{\lambda} = \frac{\nu / \mu}{\lambda / \mu} = \frac{\beta}{\rho}.$$

В результате получим:

$$P_n = \frac{\beta}{\rho} \cdot \frac{\frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{s \rho^s}{\prod_{j=1}^s (m + j\beta)}}{\sum_{k=0}^m \frac{\rho^k}{k!} + \frac{\rho^m}{m!} \sum_{s=1}^{\infty} \frac{\rho^s}{\prod_{j=1}^s (m + j\beta)}}. \quad (4.76)$$

Формула (4.76) также содержит бесконечные суммы. Для грубой оценки ошибки, происходящей от отбрасывания всех членов сумм, начиная с r -го, можно воспользоваться формулой [9]:

$$\sum_{s=r}^{\infty} \frac{s \rho^s}{\prod_{j=1}^s (m + j\beta)} < \frac{\left(\frac{\rho}{\beta}\right)^r}{(r-1)!} e^{\frac{\rho}{\beta}}. \quad (4.77)$$

По известной вероятности отказа P_n в обслуживании можно определить относительную пропускную способность Q системы, которая находится по известному соотношению:

$$Q = 1 - P_n.$$

Величина Q — это вероятность того, что требование, попавшее в систему, будет обслужено.

4.9.2. Анализ динамического режима системы М/М/т с ограниченным временем ожидания

Динамический режим это такой режим, когда вероятности состояний изменяются во времени. Динамика системы может быть описана системой обыкновенных дифференциальных уравнений. Для составления дифференциальных уравнений для вероятностей состояний системы воспользуемся размеченным графом состояний рис. 4.7 и мнемоническим правилом составления дифференциальных уравнений. Мнемоническое правило заключается в следующем [8, с.135]: производная вероятности любого состояния равна сумме потоков вероятности, переводящих систему в это состояние, минус сумма всех потоков вероятности, выводящих систему из этого состояния.

С учетом функционирования системы отметим, что число состояний и соответственно число дифференциальных уравнений будет бесконечным.

Дифференциальное уравнение относительно вероятности $P_0(t)$ нулевого состояния:

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t).$$

Дифференциальное уравнение относительно вероятности $P_1(t)$ первого состояния:

$$\frac{dP_1(t)}{dt} = \lambda P_0(t) - (\lambda + \mu)P_1(t) + 2\mu P_2(t).$$

Дифференциальное уравнение относительно вероятности $P_2(t)$ второго состояния:

$$\frac{dP_2(t)}{dt} = \lambda P_1(t) - (\lambda + 2\mu)P_2(t) + 3\mu P_3(t).$$

Далее структура уравнений повторяется. Получим k -е уравнение для $k < m$, т. е. когда отсутствует очередь и все требования, поступающие в систему, будут обслужены:

$$\frac{dP_k(t)}{dt} = \lambda P_{k-1}(t) - (\lambda + k\mu)P_k(t) + (k+1)\mu P_{k+1}(t).$$

Изменение структуры дифференциальных уравнений произойдет с m -го состояния, когда начнется учет состояний с наличием очереди.

Дифференциальное уравнение относительно вероятности $P_m(t)$ для m -го состояния:

$$\frac{dP_m(t)}{dt} = \lambda P_{m-1}(t) - (\lambda + m\mu)P_m(t) + (m\mu + \nu)P_{m+1}(t).$$

Дифференциальное уравнение для следующего состояния:

$$\frac{dP_{m+1}(t)}{dt} = \lambda P_m(t) - (\lambda + m\mu + \nu)P_{m+1}(t) + (m\mu + 2\nu)P_{m+2}(t).$$

Далее структура дифференциальных уравнений повторяется. Получим дифференциальное уравнение для $(m+s)$ -го состояния, $s \geq 1$:

$$\frac{dP_{m+s}(t)}{dt} = \lambda P_{m+s-1}(t) - (\lambda + m\mu + s\nu)P_{m+s}(t) + [m\mu + (s+1)\nu]P_{m+s+1}(t).$$

Таким образом, получаем следующую бесконечную систему дифференциальных уравнений:

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t);$$

$$\frac{dP_1(t)}{dt} = \lambda P_0(t) - (\lambda + \mu)P_1(t) + 2\mu P_2(t);$$

$$\frac{dP_2(t)}{dt} = \lambda P_1(t) - (\lambda + 2\mu)P_2(t) + 3\mu P_3(t);$$

.....

$$\frac{dP_k(t)}{dt} = \lambda P_{k-1}(t) - (\lambda + k\mu)P_k(t) + (k+1)\mu P_{k+1}(t);$$

.....

$$\frac{dP_{m-1}(t)}{dt} = \lambda P_{m-2}(t) - [\lambda + (m-1)\mu]P_{m-1}(t) + m\mu P_m(t).$$

$$\frac{dP_m(t)}{dt} = \lambda P_{m-1}(t) - (\lambda + m\mu)P_m(t) + (m\mu + \nu)P_{m+1}(t);$$

$$\frac{dP_{m+1}(t)}{dt} = \lambda P_m(t) - (\lambda + m\mu + \nu)P_{m+1}(t) + (m\mu + 2\nu)P_{m+2}(t);$$

.....

$$\frac{dP_{m+s}(t)}{dt} = \lambda P_{m+s-1}(t) - (\lambda + m\mu + s\nu)P_{m+s}(t) + [m\mu + (s+1)\nu]P_{m+s+1}(t);$$

.....

(4.78)

Систему (4.78) дополним естественными начальными условиями и нормировочным соотношением:

$$P_0(0) = 1, \quad P_i(0) = 0, \quad i \geq 1, \tag{4.79}$$

$$\sum_{k=0}^{\infty} P_k(t) = 1. \tag{4.80}$$

За счет нормировочного соотношения число дифференциальных уравнений может быть уменьшено на единицу. Если принять, что в системе ограничено не только время пребывания в очереди, но ограничена и длина очереди (конечный накопитель), то система дифференциальных уравнений может быть получена по следующему размеченному графу состояний, показанному на рис. 4.8:

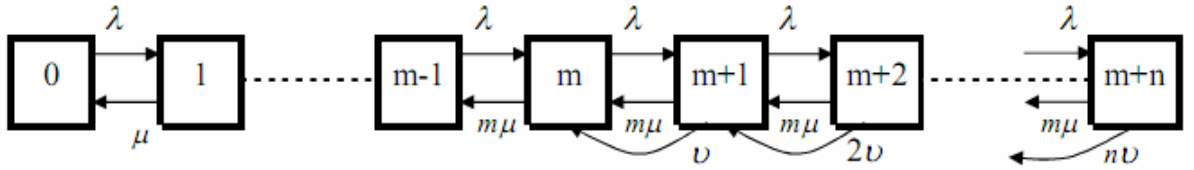


Рис. 4.8. Конечный граф состояний системы М/М/т с ограниченным временем ожидания и ограниченной длиной очереди

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t);$$

$$\frac{dP_1(t)}{dt} = \lambda P_0(t) - (\lambda + \mu)P_1(t) + 2\mu P_2(t);$$

$$\dots\dots\dots$$

$$\frac{dP_k(t)}{dt} = \lambda P_{k-1}(t) - (\lambda + k\mu)P_k + (k+1)\mu P_{k+1}(t);$$

$$\dots\dots\dots$$

$$\frac{dP_{m-1}(t)}{dt} = \lambda P_{m-2}(t) - [\lambda + (m-1)\mu]P_{m-1} + m\mu P_m(t).$$

$$\frac{dP_m(t)}{dt} = \lambda P_{m-1}(t) - (\lambda + m\mu)P_m(t) + (m\mu + \nu)P_{m+1}(t);$$

$$\frac{dP_{m+1}(t)}{dt} = \lambda P_m(t) - (\lambda + m\mu + \nu)P_{m+1}(t) + (m\mu + 2\nu)P_{m+2}(t);$$

$$\dots\dots\dots$$

$$\frac{dP_{m+s}(t)}{dt} = \lambda P_{m+s-1}(t) - (\lambda + m\mu + s\nu)P_{m+s}(t) + [m\mu + (s+1)\nu]P_{m+s+1}(t); \quad (4.81)$$

$$\dots\dots\dots$$

$$\frac{dP_{m+n}(t)}{dt} = \lambda P_{m+n-1}(t) - [m\mu + n\nu]P_{m+n}(t).$$

Систему (4.81) также можно дополнить естественными начальными условиями и нормировочным соотношением:

$$P_0(0) = 1, \quad P_i(0) = 0, \quad 1 \leq i \leq (m+n), \quad (4.82)$$

$$\sum_{k=0}^{m+n} P_k(t) = 1, \quad (4.83)$$

где n — допустимая длина очереди, $(m+n)$ — допустимое число требований в системе.

Очевидно, что если решение системы (4.81) будет стремиться к установившемуся значению, то это значит, что в системе возможен стационарный режим, определяемый конечными, предельными, или стационарными вероятностями.

4.9.3. Пример анализа динамического режима системы М/М/5/9 с ограниченным временем ожидания и ограниченной длиной очереди

Рассмотрим численное решение дифференциальных уравнений для системы М/М/5/9 с ограниченным временем ожидания в вычислительном пакете MATLAB. В соответствии с рис.4.8 и (4.81) будем иметь следующую систему дифференциальных уравнений:

$$\begin{aligned}
 \frac{dP_0(t)}{dt} &= -\lambda P_0(t) + \mu P_1(t); \\
 \frac{dP_1(t)}{dt} &= \lambda P_0(t) - (\lambda + \mu)P_1(t) + 2\mu P_2(t); \\
 \frac{dP_2(t)}{dt} &= \lambda P_1(t) - (\lambda + 2\mu)P_2 + 3\mu P_3(t); \\
 \frac{dP_3(t)}{dt} &= \lambda P_2(t) - (\lambda + 3\mu)P_3 + 4\mu P_4(t); \\
 \frac{dP_4(t)}{dt} &= \lambda P_3(t) - (\lambda + 4\mu)P_4 + 5\mu P_5(t); \\
 \frac{dP_5(t)}{dt} &= \lambda P_4(t) - (\lambda + 5\mu)P_5(t) + (5\mu + \nu)P_6(t); \\
 \frac{dP_6(t)}{dt} &= \lambda P_5(t) - (\lambda + 5\mu + \nu)P_6(t) + (5\mu + 2\nu)P_7(t); \\
 \frac{dP_7(t)}{dt} &= \lambda P_6(t) - (\lambda + 5\mu + 2\nu)P_7(t) + (5\mu + 3\nu)P_8(t); \\
 \frac{dP_8(t)}{dt} &= \lambda P_7(t) - (\lambda + 5\mu + 3\nu)P_8(t) + (5\mu + 4\nu)P_9(t); \\
 \frac{dP_9(t)}{dt} &= \lambda P_8(t) - (5\mu + 4\nu)P_9(t).
 \end{aligned} \tag{4.84}$$

В матричной форме система (4.84) будет иметь вид:

$$\frac{dP}{dt} = AP,$$

где P — искомый вектор вероятностей, размера 10×1 , A — матрица коэффициентов размера 10×10 с главной диагональю, поддиагональю, наддиагональю и остальные элементы нулевые.

Для системы (4.84) примем естественные начальные условия:

$$P_0(0) = 1, \quad P_i(0) = 0, \quad i = \overline{1, 9}.$$

Программный код решения примера:

```
function MM59;
%% Система М/М/5/9
%% с ограниченной очередью
```

```
clc,close all
format compact
L = 20;
M = 23.5;
V = 1.3;
m = 5;
A=[-L, M, zeros(1,8);
    L, -(L+M), 2*M, zeros(1,7);
    0, L, -(L+2*M), 3*M, zeros(1,6);
    zeros(1,2), L, -(L+3*M), 4*M, zeros(1,5);
    zeros(1,3), L, -(L+4*M), 5*M, zeros(1,4);
    zeros(1,4), L, -(L+5*M), (5*M+V), zeros(1,3);
    zeros(1,5), L, -(L+5*M+V), (5*M+2*V), zeros(1,2);
    zeros(1,6), L, -(L+5*M+2*V), (5*M+3*V), 0;
    zeros(1,7), L, -(L+5*M+3*V), (5*M+4*V);
    zeros(1,8), L, -(5*M+4*V)];

P0 = [1; zeros(length(A)-1,1)];
T = [0, 2];
[t,P] = ode45(@syst1,T,P0, [], A);

format long
Pycm = P(end,:)'
S = sum(P(end,:))

plot(t,P,'linew',2),
grid on,
str = sprintf('%s%s%s%s%s%s', '\bf\fontsize{12}\lambda = ',L,...
```

```

'\newline\mu = ',M,'\newline\upsilon = ', V);
text(1.5,0.9,str);
text(0.06,0.52,'\it\fontsize{12}\leftarrow P_o\rm(\itt\rm)'),
text(T(end)+0.02,0,'\bf\it\fontsize{12}t'),
text(-0.1,1.05,'\bf\it\fontsize{12} P\rm\bf(\itt\rm\bf)'),
ylabel('\bf Вероятность'); ylim([0 1.02]),
set(gcf, 'color', [0.7,0.7,0.7]),

%% M-функция описания дифференциальных уравнений
function f = syst1(t,P, A)
f = A*P;

```

Временные зависимости вероятностей системы М/М/5/9 с ограниченным временем ожидания показаны на рис. 4.9.

Результаты выполнения программы — расчет установившихся вероятностей $P_{уст}$ состояний и проверка нормировочного условия S :

$P_{уст} =$

```

0.426932915415650
0.363347125800897
0.154615939399011
0.043862363506364
0.009332852626016
0.001588194324108
0.000267596243879
0.000044460990680
0.000007360377159
0.000001191316236

```

$S =$

1.0000000000000000

Как видно, нормировочное условие выполняется.

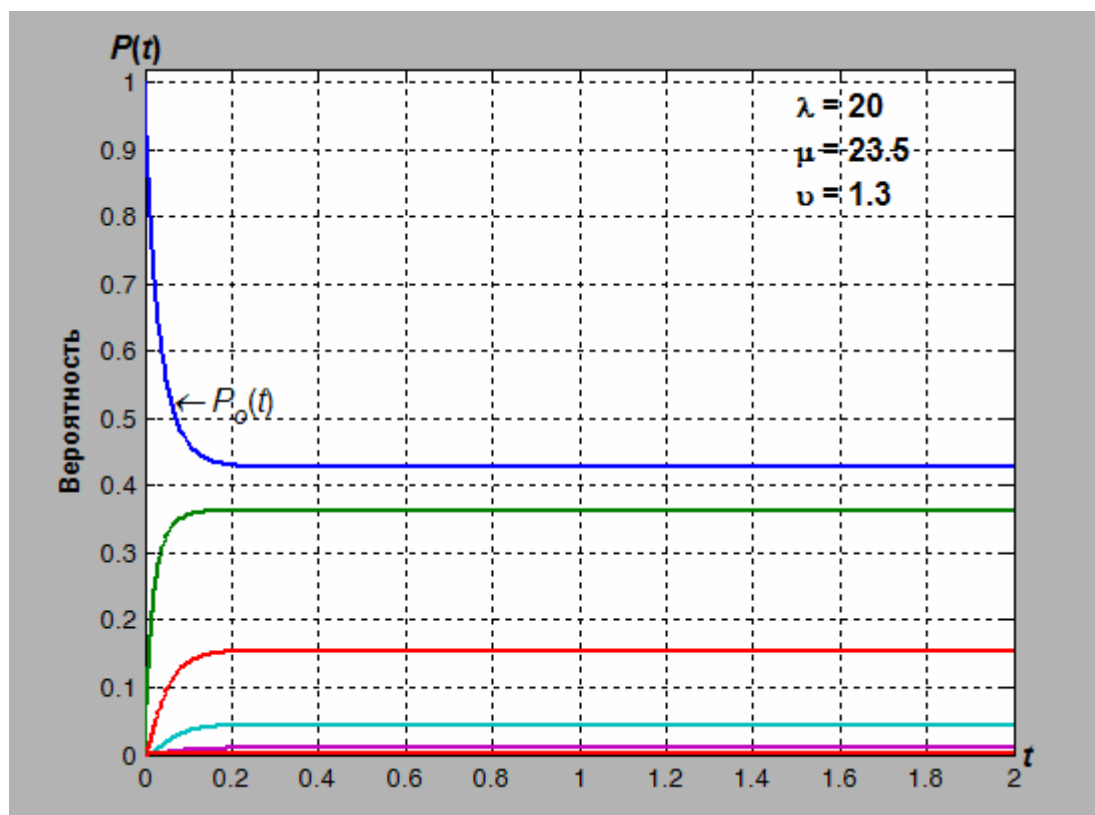


Рис. 4.9. Вероятности состояний системы М/М/5/9
с ограниченным временем ожидания

По приведенной программе можно провести исследования при различных параметрах системы, т. е. при различных λ , μ , ν . Следует отметить, что при изменении интенсивности ухода из очереди "нетерпеливых" требований стационарные значения вероятностей будут меняться только с номера, превышающего число приборов обслуживания.

Сравним полученные значения конечных вероятностей со значениями вероятностей тех же состояний, определяемых по формулам (4.69), (4.71), (4.73), найденных из условия стационарного режима данной системы с параметрами $\lambda = 20$, $\mu = 23.5$, $\nu = 1.3$.

Для расчета используем следующую MATLAB-программу, в которой бесконечное суммирование ограничено величиной, равной 150:

```
clear,clc
L = 20;
M = 23.5;
v = 1.3;
b = v/M;
r = L/M;
m = 5;
n = 9;
S01 = 0;
for k = 0:m
    S01 = S01 + r^k/factorial(k);
end
S02 = 0;
for s = 1 : 150
    k = 1:s;
    Pr = prod(m + k*b);
    S02 = S02 + r^s/Pr;
end
format long
p0 = 1/(S01+(r^m/factorial(m))*S02);
q = 0;
for k = 0 : m
    q = q + 1;
    p15(q) = p0*r^k/factorial(k);
end
w = 0;
for t = 1 : (n - m)
    w = w+1;
```

```
z = 1 : t;  
p69(w)=p0*r^(m+t)/(factorial(m)*prod(m+z*b));  
end  
Pcontrol = [p15'; p69'];  
disp('    Pcontrol: ')  
  
disp(Pcontrol)
```

Результат выполнения программы:

```
Pcontrol:  
0.426932813701671  
0.363347075490784  
0.154615776804589  
0.043862631717614  
0.009332474833535  
0.001588506354644  
0.000267425312230  
0.000044533773893  
0.000007336700806  
0.000001195876252
```

Полученные результаты практически совпадают: расчеты по прямым формулам стационарного режима и через установившиеся решения дифференциальных уравнений.

4.10. Система М/М/∞//М с конечным числом источников нагрузки и бесконечным числом обслуживающих приборов

Рассматривается система, в которую поступает конечное число требований, и предполагается, что для каждого требования найдется прибор обслуживания [4]. Поэтому в обозначении системы символ бесконечности понимается условно. Для данного обозначения системы первые две буквы означают, что в систему поступает пуассоновский поток требований, которые обслуживаются по экспоненциальному закону. Для такой модели ее параметры будут следующими:

$$\lambda_k = \begin{cases} \lambda(M - k), & 0 \leq k \leq M; \\ 0 & \text{в остальных случаях;} \end{cases} \quad (4.85)$$

$$\mu_k = k\mu, \quad k = 1, 2, 3, \dots,$$

где λ_k — интенсивность поступления в систему k -го требования, $\lambda = \text{const}$ — интенсивность поступления одного требования, μ_k — интенсивность обслуживания k -го требования, $\mu = \text{const}$ — интенсивность обслуживания одного требования каждым прибором [4, с. 125]. Интенсивность поступления требований в систему уменьшается с уменьшением числа требований перед входом системы, интенсивность обслуживания системы увеличивается с увеличением числа требований, поступивших в систему, в которой всегда предоставляется прибор обслуживания. Размеченный граф состояний системы показан на рис. 4.10.

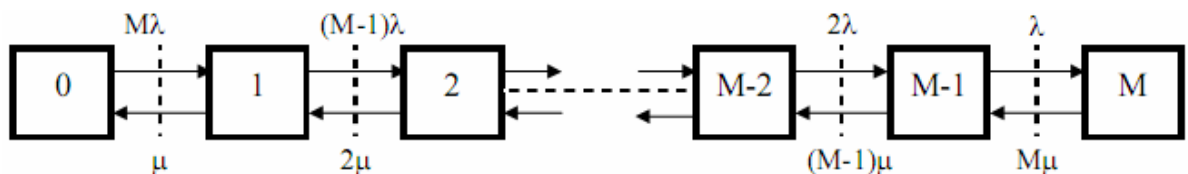


Рис. 4.10. Размеченный граф состояний системы $M/M/\infty/M$

Получим соотношения для определения стационарных вероятностей состояний. Для этого, как и прежде будем пользоваться мнемоническим правилом, предполагая, что между двумя состояниями

есть воображаемая граница (на рис. 4.10 границы показаны вертикальными штриховыми линиями).

Между нулевым и первым состояниями из равенства потоков вероятностей через воображаемую границу найдем:

$$M\lambda p_0 = \mu p_1, \Rightarrow p_1 = \frac{M\lambda}{\mu} p_0 = M\rho p_0,$$

где $\rho = \lambda / \mu$.

Между первым и вторым состояниями находим:

$$(M-1)\lambda p_1 = 2\mu p_2, \Rightarrow p_2 = \frac{(M-1)\lambda}{2\mu} p_1 = \frac{(M-1)}{2} \rho M\rho p_0 = \frac{(M-1)M\rho^2}{2} p_0.$$

Между вторым и третьим состояниями находим:

$$\begin{aligned} (M-2)\lambda p_2 &= 3\mu p_3, \Rightarrow p_3 = \frac{(M-2)\lambda}{3\mu} p_2 = \frac{(M-2)}{3} \rho \frac{(M-1)M\rho^2}{2} p_0 = \\ &= \frac{(M-2)(M-1)M\rho^3}{3!} p_0 = p_0 \frac{\rho^3}{3!} \prod_{i=0}^2 (M-i) \end{aligned}$$

или

$$\begin{aligned} p_3 &= \frac{(M-2)(M-1)M\rho^3}{3!} p_0 = p_0 \frac{\rho^3}{3!} \frac{1 \cdot 2 \cdot 3 \cdot \dots \cdot (M-3)(M-2)(M-1)M}{1 \cdot 2 \cdot 3 \cdot \dots \cdot (M-3)} = \\ &= p_0 \frac{\rho^3 M!}{3!(M-3)!} = p_0 \rho^3 C_3^M, \end{aligned}$$

где $C_3^M = \frac{M!}{3!(M-3)!}$ — биномиальные коэффициенты.

Наметившаяся закономерность в определении вероятностей состояний позволяет записать следующие выражения:

$$p_k = p_0 \rho^k \prod_{i=0}^{k-1} (M-i), \quad 0 \leq k \leq M \quad (4.86)$$

или

$$p_k = p_0 \rho^k C_k^M, \quad 0 \leq k \leq M, \quad (4.87)$$

где $C_k^M = \frac{M!}{k!(M-k)!}$ — биномиальные коэффициенты.

Из нормировочного условия определим вероятность p_0 с учетом (4.87):

$$\begin{aligned} \sum_{k=0}^M p_k = 1, \quad \Rightarrow \quad \sum_{k=0}^M p_0 \rho^k C_k^M = 1, \quad p_0 = \left[\sum_{k=0}^M \rho^k C_k^M \right]^{-1} = \\ = \left[1 \cdot C_0^M + \rho C_1^M + \rho^2 C_2^M + \dots + \rho^M C_M^M \right]^{-1} = \left[1 + \rho C_1^M + \rho^2 C_2^M + \dots + \rho^M \right]^{-1}. \end{aligned}$$

Последний ряд представляет собой бином Ньютона:

$$\sum_{k=0}^M \rho^k C_k^M = 1 + \rho C_1^M + \rho^2 C_2^M + \dots + \rho^M = (1 + \rho)^M. \quad (4.88)$$

Тогда:

$$p_0 = \frac{1}{(1 + \rho)^M}. \quad (4.89)$$

Таким образом, можем записать следующее соотношение для расчета стационарных вероятностей системы М/М/∞/М:

$$p_k = \begin{cases} \frac{\rho^k C_k^M}{(1 + \rho)^M}, & 0 \leq k \leq M; \\ 0 & \text{в остальных случаях.} \end{cases} \quad (4.90)$$

Вычислим среднее число N требований в системе с применением искусственного дифференцирования и с учетом формул (4.88), (4.90):

$$N = \sum_{k=0}^M k p_k = \sum_{k=0}^M k \frac{\rho^k C_k^M}{(1 + \rho)^M} = \frac{\rho}{(1 + \rho)^M} \frac{\partial}{\partial \rho} \sum_{k=0}^M \rho^k C_k^M = \frac{\rho}{(1 + \rho)^M} \cdot \frac{\partial}{\partial \rho} (1 + \rho)^M.$$

После дифференцирования получим:

$$N = \frac{\rho}{(1 + \rho)^M} M (1 + \rho)^{M-1} = \frac{\rho M (1 + \rho)^M (1 + \rho)^{-1}}{(1 + \rho)^M} = \frac{M \rho}{1 + \rho}, \quad (4.91)$$

где $\rho = \lambda / \mu$, M — начальное число требований в источнике нагрузки, из которого требования поступают в систему.

Отметим, что система $M/M/\infty/M$ является эргодической. Здесь это связано с условием функционирования системы: сколько бы требований не поступало в систему с любой конечной интенсивностью, всегда найдется обслуживающий прибор для вновь поступившего требования.

4.11. Система $M/M/m/K/M$ с конечным числом источников нагрузки, с m обслуживающими приборами и конечным накопителем

Система $M/M/m/K/M$ — это система с пуассоновским входящим потоком требований, с экспоненциальным законом обслуживания в m приборах, с допустимым числом требований в системе, не превышающим K , и с ограниченным числом источников нагрузки, которые создают поток из M требований. Общее число K требований в системе заключено в интервале $m \leq K \leq M$, где последняя буква M — это число требований, формируемых конечным числом источников нагрузки. Предполагается, что требования, поступающие в систему тогда, когда в ней уже имеется K требований, теряются и немедленно возвращаются в группу поступающих так, как будто бы они полностью обслужены [4]. Для описанного функционирования системы и ее заданного буквенного обозначения можно определить ее параметры в соответствии с процессом размножения и гибели в следующем виде:

$$\lambda_k = \begin{cases} \lambda(M - k), & 0 \leq k \leq K - 1, \quad \lambda = \text{const}; \\ 0 & \text{в остальных случаях;} \end{cases}$$

$$\mu_k = \begin{cases} k\mu, & 0 \leq k \leq m, \quad \mu = \text{const}; \\ m\mu, & k \geq m. \end{cases}$$

Диаграмма интенсивностей переходов для системы М/М/м/К/М будет представлять собой конечный размеченный граф состояний, который показан на рис. 4.11.

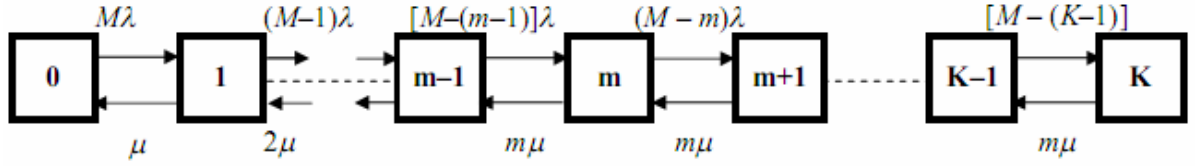


Рис. 4.11. Граф состояний системы М/М/м/К/М

На рис. 4.11 вертикальными штриховыми линиями размечены границы между состояниями системы. Определим стационарные вероятности системы по мнемоническому правилу и заданному графу состояний системы М/М/м/К/М.

Рассмотрим потоки вероятностей через границу между нулевым и первым состояниями:

$$M\lambda p_0 = \mu p_1, \Rightarrow p_1 = p_0 \frac{M\lambda}{\mu}.$$

Рассмотрим потоки вероятностей через границу между первым и вторым состояниями с учетом расчета вероятности первого состояния на предыдущем шаге:

$$(M-1)\lambda p_1 = 2\mu p_2, \Rightarrow p_2 = p_1 \frac{(M-1)\lambda}{2\mu} = p_0 \frac{(M-1)M\lambda^2}{2\mu^2}.$$

Рассмотрим потоки вероятностей через границу между вторым и третьим состояниями:

$$\begin{aligned} (M-2)\lambda p_2 &= 3\mu p_3, \Rightarrow p_3 = p_2 \frac{(M-2)\lambda}{3\mu} = p_0 \frac{(M-2)(M-1)M\lambda^3}{2 \cdot 3\mu^3} = \\ &= p_0 \rho^3 \frac{(M-2)(M-1)M}{2 \cdot 3}, \end{aligned}$$

где $\rho = \lambda / \mu$. В дальнейшем будем предполагать, что $\rho < m$.

В определении вероятностей состояний в пределах от 0-го до $(m-1)$ -го намечается определенная закономерность. Прежде чем записывать вероятность для произвольного состояния до состояния $m-1$, выполним преобразования на примере вероятности третьего состояния:

$$\begin{aligned} p_3 &= p_0 \rho^3 \frac{(M-2)(M-1)M}{1 \cdot 2 \cdot 3} = p_0 \rho^3 \frac{[1 \cdot 2 \cdot \dots \cdot (M-4)(M-3)](M-2)(M-1)M}{3! [1 \cdot 2 \cdot \dots \cdot (M-4)(M-3)]} = \\ &= p_0 \rho^3 \frac{M!}{3!(M-3)!} = p_0 \rho^3 C_M^3, \end{aligned}$$

где C_M^3 — биномиальные коэффициенты.

Таким образом, можем записать следующее выражение для определения стационарных вероятностей состояний от 0 до $m-1$:

$$p_k = p_0 \rho^k \frac{M!}{k!(M-k)!} = p_0 \rho^k C_M^k, \quad 0 \leq k \leq m-1. \quad (4.92)$$

В частности,

$$p_{m-1} = p_0 \rho^{m-1} \frac{M!}{(m-1)![M-(m-1)]!} = p_0 \rho^{m-1} C_M^{m-1} = p_0 \frac{\lambda^{m-1}}{\mu^{m-1}} C_M^{m-1}. \quad (4.93)$$

Следующая группа состояний заключена между m -м и K -м номерами. Рассмотрим поток вероятностей через границу между состояниями $(m-1)$ и m с учетом (4.93):

$$\begin{aligned} [M-(m-1)]\lambda p_{m-1} &= m\mu p_m, \Rightarrow p_m = p_{m-1} \frac{[M-(m-1)]\lambda}{m\mu} = \\ &= p_0 \frac{[M-(m-1)]\lambda^m}{m\mu^m} C_M^{m-1} = p_0 \frac{[M-(m-1)]\lambda^m}{m\mu^m} \cdot \frac{M!}{(m-1)![M-(m-1)]!} = \\ &= p_0 \frac{\lambda^m}{m\mu^m} \cdot \frac{M!}{(m-1)![M-m]!} = p_0 \frac{\lambda^m M!}{\mu^m m!(M-m)!}. \end{aligned}$$

Рассмотрим поток вероятностей через границу между состояниями m и $(m+1)$ с учетом предыдущего расчета:

$$\begin{aligned}
 (M-m)\lambda p_m &= m\mu p_{m+1}, \quad \Rightarrow \quad p_{m+1} = p_m \frac{(M-m)\lambda}{m\mu} = \\
 &= p_0 \frac{\lambda^{m+1} (M-m) M!}{\mu^{m+1} (M-m)! m! m} = p_0 \frac{\lambda^{m+1} M!}{\mu^{m+1} [M-(m+1)]! m! m} = \\
 &= p_0 \rho^{m+1} \frac{M! (m+1)!}{(m+1)! [M-(m+1)]! m! m},
 \end{aligned}$$

где $\rho = \lambda / \mu$.

Рассмотрим поток вероятностей через границу между состояниями $(m+1)$ и $(m+2)$:

$$\begin{aligned}
 [M-(m+1)]\lambda p_{m+1} &= m\mu p_{m+2}, \quad \Rightarrow \quad p_{m+2} = p_{m+1} \frac{[M-(m+1)]\lambda}{m\mu} = \\
 &= p_0 \frac{\lambda^{m+2} [M-(m+1)] M! (m+1)!}{\mu^{m+2} (m+1)! [M-(m+1)]! m! m^2} = p_0 \frac{\rho^{m+2} M! (m+1)! (m+2)}{(m+1)! (m+2) [M-(m+2)]! m! m^2} = \\
 &= p_0 \frac{\rho^{m+2} M! (m+2)!}{(m+2)! [M-(m+2)]! m! m^2} = p_0 \rho^{m+2} C_M^{m+2} \frac{(m+2)!}{m! m^2}.
 \end{aligned}$$

Теперь можно записать общее выражение для расчета вероятности произвольного состояния от m до K :

$$p_{m+r} = p_0 \rho^{m+r} C_M^{m+r} \frac{(m+r)!}{m! m^r}, \quad r = 0, 1, 2, \dots, K-m. \quad (4.94)$$

В выражении (4.94) произведем замену: $m+r=k$, $r=k-m$. Тогда

$$p_k = p_0 \rho^k C_M^k \frac{k!}{m! m^{k-m}} = p_0 \rho^k C_M^k \frac{k! m^{m-k}}{m!}, \quad m \leq k \leq K. \quad (4.95)$$

Для расчета вероятностей состояний по приведенным формулам необходимо еще определить вероятность нулевого состояния, которая может быть найдена из нормировочного условия с учетом (4.92), (4.95):

$$\sum_{k=0}^K p_k = 1, \quad \sum_{k=0}^{m-1} p_0 \rho^k C_M^k + \sum_{k=m}^K p_0 \rho^k C_M^k \frac{k! m^{m-k}}{m!} = 1,$$

$$p_0 = \left[\sum_{k=0}^{m-1} \rho^k C_M^k + \sum_{k=m}^K \rho^k C_M^k \frac{k! m^{m-k}}{m!} \right]^{-1}. \quad (4.96)$$

Рассмотрим случай, когда система М/М/м/К/М работает в режиме чистых потерь, т. е. когда параметры системы удовлетворяют условию $M \geq K = m$. В этом случае число требований в системе не может превышать числа обслуживающих приборов. Расчет вероятностей состояний будет определяться формулой (4.92) и формулами (4.95), (4.96) при $K = m$:

$$p_k = \frac{\rho^k C_M^k}{\sum_{i=0}^m \rho^i C_M^i}, \quad k = 0, 1, 2, \dots, m. \quad (4.97)$$

Распределение вероятностей в соответствии с формулой (4.97) называется распределением Энгсета [4].

4.12. Пример моделирования системы М/М/м/К/М

Для моделирования системы М/М/м/К/М получим дифференциальные уравнения относительно вероятностей состояний в соответствии с мнемоническим правилом составления уравнений Колмогорова. Примем следующие параметры системы:

$M = 15$ — число источников нагрузки,

$m = 5$ — число обслуживающих приборов,

$K = 9$ — допустимое число требований в системе,

$\lambda = 2 \text{ с}^{-1}$ — интенсивность поступления одного требования,

$\mu = 7 \text{ с}^{-1}$ — интенсивность обслуживания требования одним прибором.

Тогда система принимает следующее обозначение: М/М/5/9/15. Число дифференциальных уравнений, описывающих поведение системы,

будет равно 10, соответствующее допустимому числу требований в системе и плюс состояние, соответствующее отсутствию требований в системе.

Для составления уравнений используем размеченный граф состояний системы, показанный на рис. 4.11. Воспользуемся также выводом вероятностей состояний, принимая во внимание, что алгебраическая сумма потоков вероятностей будет равна производной от вероятности рассматриваемого состояния, которое охвачено воображаемой поверхностью.

Рассмотрим нулевое состояние. Из него выходит поток вероятности и входит поток вероятности. Первый поток считаем отрицательным, второй — положительным. Дифференциальное уравнение будет иметь вид:

$$\frac{dP_0(t)}{dt} = -M\lambda P_0(t) + \mu P_1(t).$$

Рассмотрим первое состояние:

$$\frac{dP_1(t)}{dt} = M\lambda P_0(t) - [(M-1)\lambda + \mu]P_1(t) + 2\mu P_2(t).$$

Рассмотрим второе состояние:

$$\frac{dP_2(t)}{dt} = (M-1)\lambda P_1(t) - [(M-2)\lambda + 2\mu]P_2(t) + 3\mu P_3(t).$$

Рассмотрим третье состояние:

$$\frac{dP_3(t)}{dt} = (M-2)\lambda P_2(t) - [(M-3)\lambda + 3\mu]P_3(t) + 4\mu P_4(t).$$

Рассмотрим четвертое состояние:

$$\frac{dP_4(t)}{dt} = (M-3)\lambda P_3(t) - [(M-4)\lambda + 4\mu]P_4(t) + 5\mu P_5(t).$$

Рассмотрим пятое состояние:

$$\frac{dP_5(t)}{dt} = (M - 4)\lambda P_4(t) - [(M - 5)\lambda + 5\mu]P_5(t) + 5\mu P_6(t).$$

Начиная с пятого состояния и до 9-го, интенсивность обслуживания остается неизменной и равной 5μ . Рассмотрим шестое состояние:

$$\frac{dP_6(t)}{dt} = (M - 5)\lambda P_5(t) - [(M - 6)\lambda + 5\mu]P_6(t) + 5\mu P_7(t).$$

Рассмотрим седьмое состояние:

$$\frac{dP_7(t)}{dt} = (M - 6)\lambda P_6(t) - [(M - 7)\lambda + 5\mu]P_7(t) + 5\mu P_8(t).$$

Рассмотрим восьмое состояние:

$$\frac{dP_8(t)}{dt} = (M - 7)\lambda P_7(t) - [(M - 8)\lambda + 5\mu]P_8(t) + 5\mu P_9(t).$$

Рассмотрим последнее девятое состояние:

$$\frac{dP_9(t)}{dt} = (M - 8)\lambda P_8(t) - 5\mu P_9(t).$$

Полученная система уравнений является однородной линейной системой обыкновенных дифференциальных уравнений с постоянными коэффициентами 10-го порядка. Для решения начальной задачи зададим естественные начальные условия:

$$P_0(0) = 1, \quad P_i(0) = 0, \quad i = \overline{1, 9}. \quad (4.98)$$

Представим систему дифференциальных уравнений в матричном виде:

$$\frac{dP(t)}{dt} = AP(t), \quad (4.99)$$

где $P(t)$ — вектор вероятностей состояний размера 10×1 , A — матрица постоянных коэффициентов размера 10×10 , имеющая ленточно-диагональную форму с учетом значений параметров $M = 15$, $\lambda = 2$, $\mu = 7$ следующий вид:

$$A = \begin{bmatrix} -30 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 30 & -35 & 14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 28 & -40 & 21 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 26 & -45 & 28 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 24 & -50 & 35 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 22 & -55 & 35 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 & -53 & 35 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 18 & -51 & 35 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16 & -49 & 35 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 14 & -35 \end{bmatrix}. \quad (4.100)$$

Решение векторного дифференциального уравнения (4.99) с начальными условиями (4.98) и заданной матрицей коэффициентов (4.100) выполним численным методом в системе MATLAB по приводимой ниже программе:

```
function MM_5_9_15
clc,close all
L = 2;
M = 7;
MM = 15;

A=[-MM*L, M zeros(1,8);
MM*L, -((MM-1)*L+M), 2*M zeros(1,7);
0, (MM-1)*L, -((MM-2)*L+2*M), 3*M, zeros(1,6);
0, 0, (MM-2)*L, -((MM-3)*L+3*M), 4*M, zeros(1,5);
zeros(1,3), (MM-3)*L, -((MM-4)*L+4*M), 5*M, zeros(1,4);
zeros(1,4), (MM-4)*L, -((MM-5)*L+5*M), 5*M zeros(1,3);
zeros(1,5), (MM-5)*L, -((MM-6)*L+5*M), 5*M, 0, 0;
```

```
zeros(1,6), (MM-6)*L, -((MM-7)*L+5*M), 5*M, 0;
zeros(1,7), (MM-7)*L, -((MM-8)*L+5*M), 5*M;
zeros(1,8), (MM-8)*L, -5*M];
P0 = [1;zeros(length(A)-1,1)];
T = [0, 2];
[t,p] = ode45(@prob412,T,P0, [], A);

Pycm=p(end,:)', S = sum(Pycm)
plot(t,p, 'linew',2),
grid on,
xlabel('\bf\it - - - - - t - - - - -'),
ylabel('\bf Вероятность'),
str = sprintf('%s%s%s%s',
'\bf\it\fontsize{14}\lambda\rm\bf\fontsize{12} = ', L,...
'\it\fontsize{14} \mu\rm\bf\fontsize{12} = ', M);
text(1.25,.75, str),
text(-0.1,1.06,'\bf\it\fontsize{12} P\rm\bf(\it\rm\bf)'),
ylim([0, 1.02]), set(gcf,'color',[0.7,0.7,0.7])

%% М-функция описания правых частей дифференциальных
уравнения
function f = prob412(t,p, A)
f = A*p;
```

Выполнение программы приводит к следующим результатам и диаграмме изменения вероятностей (рис. 4.12):

Pycm =

0.022072220530025

0.094593943720601
 0.189193416160514
 0.234224594900032
 0.200788369594891
 0.126183505449716
 0.072124836995376
 0.037081824327135
 0.016956070431645
 0.006781217890065
 S =
 1.0000000000000000

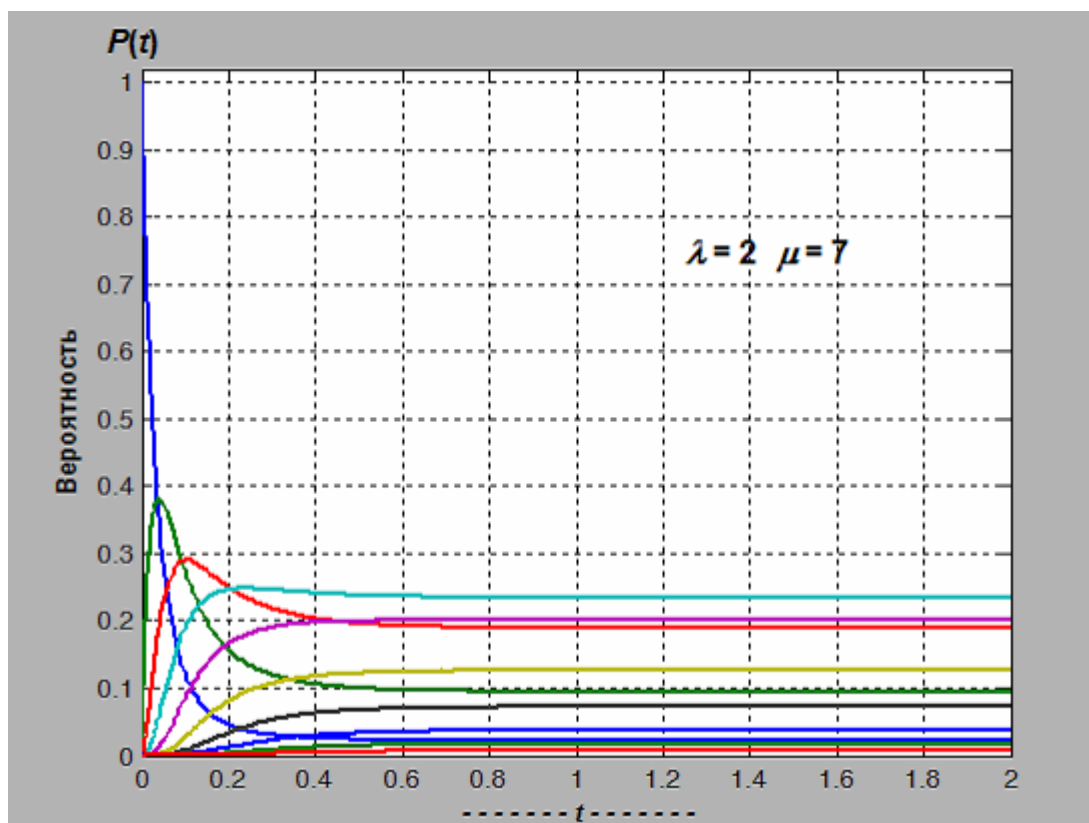


Рис. 4.12. Временные зависимости вероятностей системы М/М/5/9/15

Для проверки значений стационарных вероятностей, полученных на основе решения дифференциальных уравнений (установившихся

решений), рассчитаем эти же вероятности по формулам (4.92), (4.96), (4.96). Для этого используем следующую MATLAB-программу:

```
clear,clc
L = 2;
M = 7;
r = L/M;
m = 5;
K = 9;
MM = 15;
S01 = 0;
for k = 0 : m-1
    S01 = S01+(r^k*factorial(MM)/(factorial(k)*factorial(MM-k)));
end
S02 = 0;
for k = m : K
    CK15 = factorial(MM)/(factorial(k)*factorial(MM-k));
    S02 = S02+(r^k*CK15*factorial(k)*m^(m-k))/factorial(m);
end
p0 = 1/(S01 + S02);
q = 0;
for k = 0 : m-1
    q = q + 1;
    p1m(q) = p0*r^k*factorial(MM)/(factorial(k)*factorial(MM-k));
end
w = 0;
for k = m : K
    w = w+1;
    pmK(w) = p0*r^k*factorial(MM)/(factorial(k)*factorial(MM-
```

```
k))*factorial(k)*m^(m-k)/factorial(m);  
end  
Pcm =[p1m';pmK']
```

Результат выполнения программы:

Pcm =

```
0.022072090921605  
0.094594675378308  
0.189189350756615  
0.234234434270095  
0.200772372231510  
0.126199776831235  
0.072114158189277  
0.037087281354485  
0.016954185762050  
0.006781674304820
```

Как видно, результаты расчетов одних и тех же вероятностей ($P_{уст}$ и P_{cm}) практически совпадают, как при расчете по решению дифференциальных уравнений заданной системы, так и по расчетам конечных формул, выведенных из условия стационарного режима заданной системы массового обслуживания.

ГЛАВА 5. МНОГОФАЗНЫЕ СИСТЕМЫ ОБСЛУЖИВАНИЯ

В этой главе рассматриваются двух и трехфазные модели массового обслуживания с нулевой вместимостью блока ожидания, которые используются в инженерных приложениях.

5.1. Двухфазная модель с нулевой вместимостью блока ожидания

Схема рассматриваемой модели показана на рис. 5.1. Нулевая вместимость блока ожидания означает, что очереди перед каждой фазой обслуживания отсутствуют.

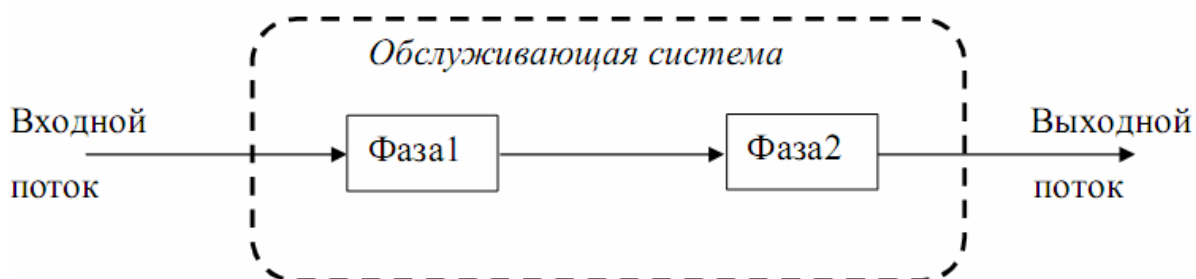


Рис. 5.1. Двухфазная модель системы обслуживания

Дальнейшее изложение в основном будет следовать материалу из [3]. Для рассматриваемой модели обслуживания входной поток предполагается пуассоновским со средней интенсивностью, равной λ . Обслуживание в каждой фазе осуществляется по экспоненциальному закону с одной и той же средней скоростью обслуживания μ . Каждое прибывающее требование должно сначала пройти через первую фазу обслуживания, а затем через вторую. При этом принимается, что очереди перед каждой фазой не должны быть.

Рассмотрим возможные состояния системы, показанной на рис. 5.1. Каждая фаза может быть занята на обслуживание, либо свободна. Поскольку очередь не допускается перед фазой, то будем говорить, что первая фаза (Фаза1 на рис. 5.1) обслуживания заблокирована, если обслуживание требования в данной фазе завершено, а вторая фаза (Фаза2 на рис. 5.1) не готова к приему данного требования по той причине, что в ней еще не закончена процедура обслуживания. Требование не имеет права на ожидание в промежутке между фазами 1 и 2. Введем обозначения 0, 1 и b , которые представляют состояния "фаза свободна", "фаза занята" и "фаза заблокирована". Состояние первой фазы будем обозначать символом i , состояние второй фазы — символом j . Множество состояний обслуживающей двухфазной системы будет следующим:

$$\{(i, j)\} = \{(0, 0), (1, 0), (0, 1), (1, 1), (b, 1)\}. \quad (5.1)$$

Далее обозначим через $p_{ij}(t)$ вероятность того, что в момент времени t система находится в состоянии (i, j) , т.е. в одном из возможных состояний в соответствии с (5.1).

Для пуассоновского потока с параметром λ вероятность отсутствия требований на интервале Δt определяется выражением: $p_0(\Delta t) = e^{-\lambda \Delta t}$. Используя разложение экспоненты в ряд, и ограничиваясь только линейными членами в силу малости Δt , будем иметь

$$p_0(\Delta t) = e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t. \quad (5.2)$$

В силу ординарности пуассоновского потока получим из условия полной группы событий значение вероятности $p_1(\Delta t)$ появления одного требования на малом интервале Δt :

$$p_0(\Delta t) + p_1(\Delta t) \approx 1, \Rightarrow p_1(\Delta t) = 1 - p_0(\Delta t),$$

$$p_1(\Delta t) = 1 - (1 - \lambda \Delta t) = \lambda \Delta t. \quad (5.3)$$

Рассмотрим вероятности переходов из одного состояния в другое на временном интервале $[t, t + \Delta t]$, где Δt — бесконечно малое положительное приращение времени. Для удобства построим таблицу 5.1, в которой левый столбец будет представлять исходные состояния системы в момент времени t , а в верхней строке будут указаны состояния, в которые за время Δt система может перейти или не перейти.

Таблица 5.1

Состояния системы в момент времени t	Состояния системы на временном интервале $[t + \Delta t]$				
	(0, 0)	(0, 1)	(1, 0)	(1, 1)	(b, 1)
	Вероятности состояний на момент времени $[t + \Delta t]$				
(0, 0)	$1 - \lambda \Delta t$	—	$\lambda \Delta t$	—	—
(0, 1)	$\mu \Delta t (1 - \lambda \Delta t)$	$1 - \mu \Delta t - \lambda \Delta t$	—	$\lambda \Delta t (1 - \mu \Delta t)$	—
(1, 0)	—	$\mu \Delta t (1 - \lambda \Delta t)$	$1 - \mu \Delta t$	—	—
(1, 1)	—	$\mu \Delta t (1 - \lambda \Delta t)$	$\mu \Delta t$	$(1 - \mu \Delta t)(1 - \mu \Delta t)$	$\mu \Delta t$
(b, 1)	—	$\mu \Delta t (1 - \lambda \Delta t)$	—	—	$1 - \mu \Delta t$

Сделаем пояснения к табл. 5.1.

Вероятность перехода системы из состояния (0,0) в то же самое состояние (0,0) за время Δt равна $1 - \lambda \Delta t$, что соответствует отсутствию требований (см. формулу 5.2). Если требований не поступило, то и состояние системы не изменилось. Если в систему поступает одно требование, то оно будет сначала на обслуживании в первой фазе, поэтому за время Δt система может оказаться в состоянии (1,0), т. е. когда в первой фазе находится одно требование, а во второй фазе требования

нет. Вероятность состояния $(1,0)$ равна $\lambda \Delta t$, что соответствует вероятности поступления одного требования (см. формулу 5.3). Переход из состояния $(0,0)$ в состояние $(0,1)$ невозможен, поскольку каждое требование последовательно проходит обслуживание в первой фазе и затем во второй. Переход из состояния $(0,0)$ в состояние $(1,1)$ за время Δt также не возможен, так как в силу ординарности пуассоновского потока за малое время Δt может появиться не более одного требования, а состояние $(1,1)$ предполагает, что за малое время Δt в системе стало два требования — одно в первой фазе, другое во второй фазе. Переход из состояния $(0,0)$ в состояние $(b,1)$ за малое время Δt невозможен, так как символ b означает заблокированное требование в первой фазе, а единица — наличие требования во второй фазе, что невозможно в силу ординарности пуассоновского потока.

Из состояния $(0,1)$ возможен переход в состояние $(0,0)$ за время Δt с вероятностью $\mu \Delta t(1 - \lambda \Delta t)$, которая означает, что за время Δt с вероятностью $\mu \Delta t$ произошло одно обслуживание требования во второй фазе и с вероятностью $1 - \lambda \Delta t$ не поступило больше ни одного требования в систему. Переход системы из состояния $(0,1)$ в то же самое состояние $(0,1)$ происходит с вероятностью $1 - \mu \Delta t - \lambda \Delta t$, которая означает, что не произошло ни одного обслуживания и не поступило ни одного требования. Переход системы из состояния $(0,1)$ в состояние $(1,0)$ за время Δt не возможен, поскольку за малый интервал времени может произойти только одно событие (в силу ординарности потока): либо вторая фаза освободится и требование покинет систему, либо поступит новое требование. А переход из состояния $(0,1)$ в состояние $(1,0)$ означает, что одно требование должно покинуть систему, а другое требование должно поступить в систему. Переход из состояния $(0,1)$ в

состояние $(1,1)$ за время Δt происходит с вероятностью $\lambda \Delta t (1 - \mu \Delta t)$, которая означает, что в систему в первую фазу поступило одно требование с вероятностью $\lambda \Delta t$ и за время Δt не произошло обслуживания требования во второй фазе, что соответствует вероятности $(1 - \mu \Delta t)$. Вероятность $\lambda \Delta t (1 - \mu \Delta t)$ — это вероятность, получаемая по теореме произведения вероятностей $\lambda \Delta t$ и $(1 - \mu \Delta t)$. Переход системы из состояния $(0,1)$ в состояние $(b,1)$ за время Δt не возможен, так как блокировка (символ b) первой фазы не может происходить, если в исходном состоянии там не было требования.

Из состояния $(1,0)$ в состояние $(0,0)$ система не может перейти за время Δt , так как требование, находящееся в первой фазе (символ 1) не может покинуть систему, т. е. система не может стать свободной от требований (состояние $0,0$), поскольку требование после обслуживания в первой фазе должно перейти на обработку во вторую фазу. Переход системы за время Δt из состояния $(1,0)$ в состояние $(0,1)$ происходит с вероятностью $\mu \Delta t (1 - \lambda \Delta t)$, где $\mu \Delta t$ означает вероятность обслуживания одного требования (в первой фазе), $(1 - \lambda \Delta t)$ означает вероятность того, что требование в систему не поступит. Переход системы из состояния $(1,0)$ в то же самое состояние происходит с вероятностью $(1 - \mu \Delta t)$, что соответствует вероятности не обслуживания требования в первой фазе за малое время Δt . Переход системы из состояния $(1,0)$ в состояние $(1,1)$ за малое время Δt не возможен. Невозможность указанного перехода объясняется тем, что практически в систему поступает два требования — одно новое в первую фазу, а второе требование из первой фазы поступает на обслуживание во вторую фазу, что в силу ординарности потока требований такое не может быть. Переход системы из состояния $(1,0)$ в состояние $(b,1)$ за малое время Δt не возможен, поскольку блокировка

первой фазы происходит при наличии в ней требования, а это будет практически означать, что в систему за малое время поступает два требования — одно новое в первую фазу, другое поступает из первой фазы во вторую. Опять в силу ординарности потока требований этого не может быть.

Переход системы из состояния $(1,1)$ в состояние $(0,0)$ за время Δt не возможен, поскольку за малый промежуток времени два требования не могут покинуть систему, так как требование из первой фазы должно пройти обслуживание во второй фазе, что уже обуславливает невозможность указанного перехода. Переход системы из состояния $(1,1)$ в состояние $(0,1)$ за время Δt происходит с вероятностью $\mu\Delta t(1 - \lambda\Delta t)$, где $\mu\Delta t$ означает вероятность обслуживания одного требования двухфазной системой, а $(1 - \lambda\Delta t)$ означает вероятность того, что в систему не поступит ни одного требования. Переход системы из состояния $(1,1)$ в состояние $(1, 0)$ за время Δt происходит с вероятностью $\mu\Delta t$, которая означает обслуживание одного требования двухфазной системой (вторая фаза стала свободной). Переход системы из состояния $(1,1)$ в то же самое состояние $(1,1)$ за время $\mu\Delta t$ происходит с вероятностью $(1 - \mu\Delta t)(1 - \mu\Delta t)$, которая означает, что требование не было обслужено ни в первой фазе, ни во второй. Переход системы из состояния $(1,1)$ в состояние $(b,1)$ за время Δt происходит с вероятностью $\mu\Delta t$, которая означает, что требование в первой фазе было обслужено, и оно заблокировалось в первой же фазе, так как очередь перед второй фазой по условию работы рассматриваемой двухфазной системой не допускается.

Переход системы из состояния $(b,1)$ в состояние $(0,0)$ за время Δt не возможен, так как если освободиться вторая фаза, то заблокированное требование в первой фазе поступит на обслуживание во вторую фазу.

Переход системы из состояния $(b,1)$ в состояние $(0,1)$ за время Δt происходит с вероятностью $\mu\Delta t(1-\lambda\Delta t)$, которая означает, что система с вероятностью $\mu\Delta t$ обслужила одно требование, а с вероятностью $(1-\lambda\Delta t)$ в систему не поступит ни одного требования, при этом заблокированное требование в первой фазе перейдет на обслуживание во вторую фазу. Переход системы из состояния $(b,1)$ в состояние $(1,0)$ за время Δt не возможен, так как если бы было обслужено требование во второй фазе и покинуло систему, то в нее поступило бы на обслуживание заблокированное требование из первой фазы. Переход системы из состояния $(b,1)$ в состояние $(1,1)$ не возможен. Состояние $(1,1)$ означает, что и в первой фазе и во второй происходит обслуживание требований, поэтому оно может быть получено, если требование из второй фазы как обслуженное покинет систему, заблокированное требование в первой фазе перейдет во вторую фазу, а в первую фазу поступит новое требование, что за малый промежуток времени Δt невозможно в силу ординарности потока требований. Переход системы из состояния $(b,1)$ в то же самое состояние $(b,1)$ происходит с вероятностью $(1-\mu\Delta t)$, которая означает, что не произошло обслуживания одного требования.

Рассмотрим теперь вероятности всех возможных состояний на момент времени $[t + \Delta t]$, пренебрегая малыми более высокого порядка чем Δt .

Рассмотрим состояние $(0,0)$ отсутствия требований в системе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{00}(t + \Delta t)$. В это состояние можно перейти двумя способами:

- если в систему не поступит ни одного требования: переход из состояния $(0,0)$ в состояние $(0,0)$,

— если в системе было одно требование во второй фазе, это требование было обслужено, и оно покинуло систему: переход из состояния $(0,1)$ в состояние $(0,0)$.

Рассмотрим первый способ перехода в состояние $(0,0)$. Вероятность отсутствия требований на момент времени t обозначим $p_{00}(t)$. Вероятность того, что требование не поступило за время Δt определяется по формуле (5.2) как $(1 - \lambda \Delta t)$, которую можно назвать переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.2) с учетом вероятности $p_{00}(t)$ и переходной вероятности $(1 - \lambda \Delta t)$:

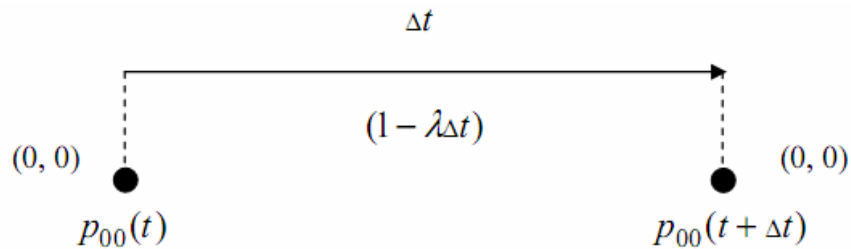


Рис. 5.2. Переход из состояния $(0, 0)$ в состояние $(0, 0)$

Из состояния $(0,0)$, определяемое вероятностью $p_{00}(t)$, за малое время Δt можно перейти в то же самое состояние $(0,0)$, определяемое вероятностью $p_{00}(t + \Delta t)$, если переходная вероятность будет равна $(1 - \lambda \Delta t)$. В данном случае переходная вероятность соответствует отсутствию поступивших требований в систему. Таким образом, искомую вероятность $p_{00}(t + \Delta t)$ требуется определить из условия, что на момент времени t вероятность была $p_{00}(t)$, а за малый промежуток времени Δt вероятность не поступления требований равна $(1 - \lambda \Delta t)$. Вероятность $p_{00}(t + \Delta t)$ будет рассчитываться по теореме умножения вероятностей, когда расчет ведется относительно условной вероятности перехода $(1 - \lambda \Delta t)$, по отношению к заданной исходной вероятности $p_{00}(t)$,

относительно которой осуществляется переход к вероятности $p_{00}(t + \Delta t)$. Поскольку поток требований по условию обладает свойством отсутствия последствия, то условная вероятность будет равна своей безусловной вероятности. В связи с этим, применяя теорему умножения вероятностей, можем записать для рассматриваемого перехода в соответствии с рис. 5.2:

$$p_{00}(t + \Delta t) = p_{00}(t)(1 - \lambda \Delta t) \quad (5.2)$$

Рассмотрим второй способ перехода в состояние $(0,0)$. Для этого также выполним графические пояснения второго способа перехода (рис. 5.3) с учетом вероятности $p_{01}(t)$ и переходной вероятности $\mu \Delta t$:

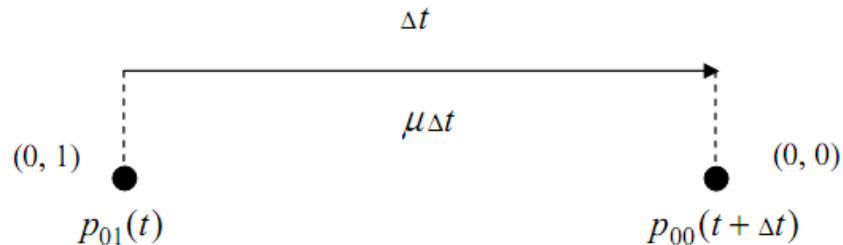


Рис. 5.3. Переход из состояния $(0, 1)$ в состояние $(0, 0)$

Из состояния $(0,1)$, определяемое вероятностью $p_{01}(t)$ на момент времени t , за малое время Δt можно перейти в состояние $(0,0)$, определяемое вероятностью $p_{00}(t + \Delta t)$ на момент времени $t + \Delta t$, если переходная вероятность будет равна $\mu \Delta t$. В данном случае переходная вероятность соответствует обслуживанию одного требования во второй фазе. Для этого условия перехода также справедлива теорема умножения вероятностей, когда условная вероятность перехода равна своей безусловной вероятности:

$$p_{00}(t + \Delta t) = p_{01}(t)\mu \Delta t \quad (5.3)$$

Теперь отметим, что два указанных перехода (рис. 5.2 и рис. 5.3) не совместны друг с другом, поэтому здесь становится справедлива теорема сложения несовместных вероятностей. А так как каждая несовместная вероятность представляет собой произведение вероятностей, то можно

записать искомую вероятность по формуле полной вероятности, которая объединяет сложение и умножение вероятностей:

$$p_{00}(t + \Delta t) = p_{00}(t)(1 - \lambda\Delta t) + p_{01}(t)\mu\Delta t. \quad (5.4)$$

Рассмотрим состояние $(0, 1)$ нахождения одного требования во второй фазе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{01}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если в систему не поступит ни одного требования и требование не будет обслужено во второй фазе: формальный переход из состояния $(0, 1)$ в то же самое состояние $(0, 1)$.
- если в первой фазе будет обслужено требование, и оно перейдет во вторую фазу: переход из состояния $(1, 0)$ в состояние $(0, 1)$.
- если систему покинет требование за счет обслуживания во второй фазе, а заблокированное требование в первой фазе перейдет во вторую фазу: переход из состояния $(b, 1)$ в состояние $(0, 1)$.

Рассмотрим первый способ перехода в состояние $(0, 1)$. Вероятность состояния $(0, 1)$ на момент времени t обозначим $p_{01}(t)$. Вероятность того, что за время Δt не поступило требований и не произошло обслуживания, определяется как $1 - \mu\Delta t - \lambda\Delta t$, которую можно назвать переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.4) с учетом вероятности $p_{01}(t)$ и переходной вероятности $1 - \mu\Delta t - \lambda\Delta t$:

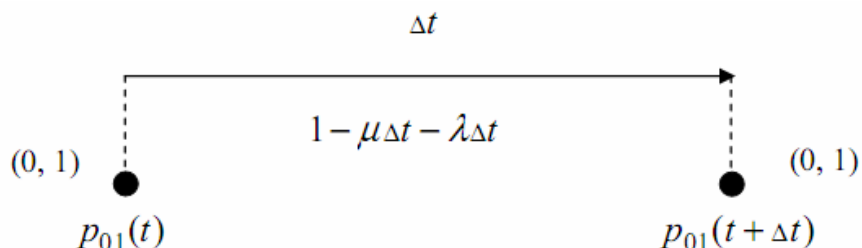


Рис. 5.4. Переход из состояния (0, 1) в состояние (0, 1)

Структура рассматриваемого перехода (рис. 5.3) аналогично ранее приведенным (рис. 5.2, рис. 5.3). Поэтому сразу запишем вероятность состояния (0, 1) на момент времени $(t + \Delta t)$, как произведение вероятностей $p_{01}(t)$ и $1 - \mu\Delta t - \lambda\Delta t$:

$$p_{01}(t + \Delta t) = p_{01}(t)(1 - \mu\Delta t - \lambda\Delta t) \quad (5.5)$$

Рассмотрим второй способ перехода в состояние (0, 1). Вероятность состояния (1, 0) на момент времени t обозначим $p_{10}(t)$. Вероятность того, что за время Δt требование будет обслужено в первой фазе, равна $\mu\Delta t$, которую можно назвать переходной вероятностью. Выполним графические пояснения (рис. 5.5) второго способа перехода с учетом вероятности $p_{01}(t)$ и переходной вероятности $\mu\Delta t$:

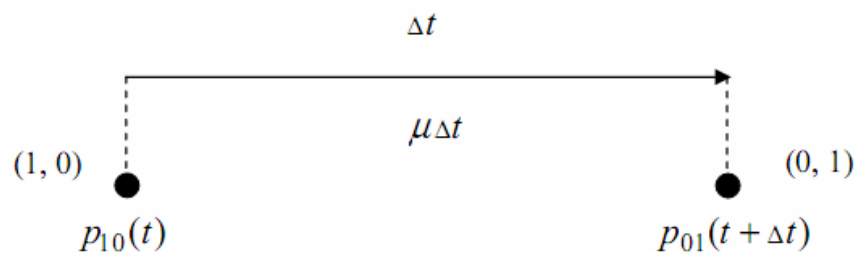


Рис. 5.5. Переход из состояния (1, 0) в состояние (0, 1)

Как и ранее определим вероятность $p_{01}(t + \Delta t)$ в соответствии с рис. 5.5 как произведение вероятностей:

$$p_{01}(t + \Delta t) = p_{10}(t)\mu\Delta t. \quad (5.6)$$

Рассмотрим третий способ перехода в состояние (0, 1). Вероятность состояния $(b, 1)$ в момент времени t обозначим как $p_{b1}(t)$. Вероятность того, что за время Δt требование будет обслужено во второй фазе, равна $\mu\Delta t$, которую также как и ранее в аналогичном случае назовем переходной вероятностью. Выполним графические пояснения (рис. 5.6)

третьего способа перехода с учетом вероятности $p_{b1}(t)$ и переходной вероятности $\mu\Delta t$:

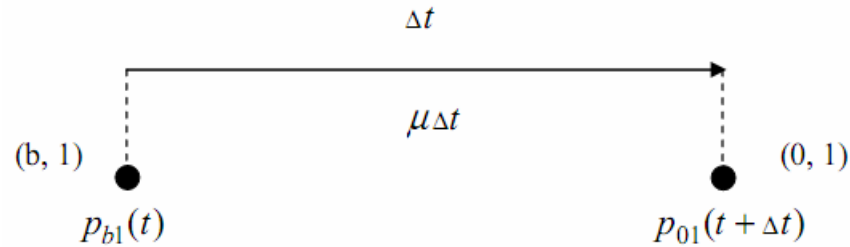


Рис. 5.6. Переход из состояния $(b, 1)$ в состояние $(0, 1)$

В соответствии с рис. 5.6 вероятность $p_{01}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{01}(t + \Delta t) = p_{b1}(t)\mu\Delta t. \quad (5.7)$$

Таким образом, было рассмотрено три взаимоисключающих способа перехода в состояние $(0, 1)$. Это дает право применить теорему сложения несовместных вероятностей или применить формулу полной вероятности с учетом (5.5), (5.6), (5.7):

$$p_{01}(t + \Delta t) = p_{01}(t)(1 - \mu\Delta t - \lambda\Delta t) + p_{10}(t)\mu\Delta t + p_{b1}(t)\mu\Delta t. \quad (5.8)$$

Рассмотрим состояние $(1, 0)$ нахождения одного требования в фазе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{10}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если из состояния $(0, 0)$ в систему за время Δt поступит одно требование: переход из состояния $(0, 0)$ в состояние $(1, 0)$,
- если в состоянии $(1, 0)$ за время Δt не произойдет обслуживания требования в первой фазе: формальный переход из состояния $(1, 0)$ в то же самое состояние $(1, 0)$,
- если в состоянии $(1, 1)$ за время Δt произойдет обслуживание требования во второй фазе: переход из состояния $(1, 1)$ в состояние $(1, 0)$.

Рассмотрим первый способ перехода в состояние $(1, 0)$. Вероятность состояния $(0, 0)$ на момент времени t обозначим $p_{00}(t)$. Вероятность того, что за время Δt поступит одно требование, равно $\lambda \Delta t$, которую можно назвать переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.7) с учетом вероятности $p_{00}(t)$ и переходной вероятности $\lambda \Delta t$:

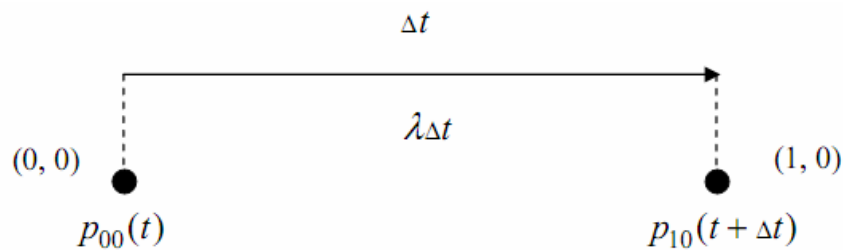


Рис. 5.7. Переход из состояния $(0, 0)$ в состояние $(1, 0)$

В соответствии с рис. 5.7 $p_{10}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{10}(t + \Delta t) = p_{00}(t) \lambda \Delta t. \quad (5.9)$$

Рассмотрим второй способ перехода в состояние $(1, 0)$. Вероятность состояния $(1, 0)$ на момент времени t обозначим $p_{10}(t)$. Вероятность того, что за время Δt не произойдет обслуживания требования в первой фазе, будет равна $1 - \mu \Delta t$, которую назовем переходной вероятностью. Графические пояснения второго способа перехода с учетом $p_{10}(t)$ и переходной вероятности $1 - \mu \Delta t$ показаны на рис. 5.8:

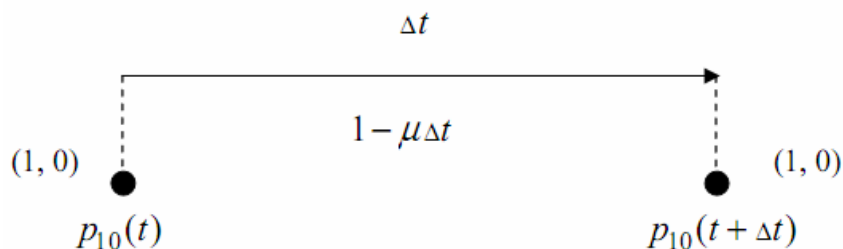


Рис. 5.8. Переход из состояния $(1, 0)$ в состояние $(1, 0)$

В соответствии с рис. 5.8 $p_{10}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{10}(t + \Delta t) = p_{10}(t)(1 - \mu\Delta t). \quad (5.10)$$

Рассмотрим третий способ перехода в состояние $(1, 0)$. Вероятность состояния $(1, 1)$ на момент времени t обозначим $p_{11}(t)$. Вероятность того, что за время Δt произойдет обслуживание требования во второй фазе, будет равна $\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.9) с учетом $p_{11}(t)$ и переходной вероятности $\mu\Delta t$:

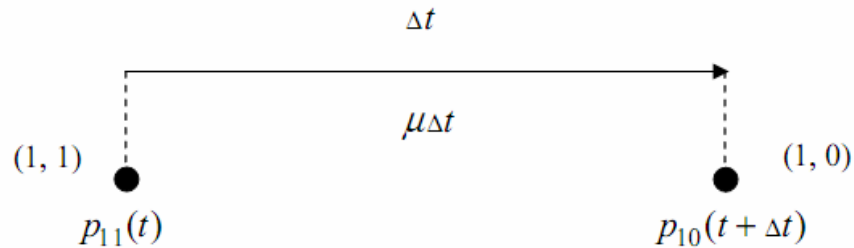


Рис. 5.9. Переход из состояния $(1, 1)$ в состояние $(1, 0)$

В соответствии с рис. 5.9 вероятность $p_{10}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{10}(t + \Delta t) = p_{11}(t)\mu\Delta t. \quad (5.11)$$

Для определения вероятности $p_{10}(t + \Delta t)$ с учетом (5.9), (5.10), (5.11) применим формулу полной вероятности:

$$p_{10}(t + \Delta t) = p_{00}(t)\lambda\Delta t + p_{10}(t)(1 - \mu\Delta t) + p_{11}(t)\mu\Delta t. \quad (5.12)$$

Рассмотрим состояние $(1, 1)$ нахождения одного требования в первой фазе и одного — во второй фазе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{11}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если в систему, находящуюся в состоянии $(0, 1)$, поступит одно требование (в первую фазу): переход из состояния $(0, 1)$ в состояние $(1, 1)$,
- если в системе не произойдет обслуживания требований ни в первой фазе, ни во второй фазе: формальный переход из состояния $(1, 1)$ в состояние $(1, 1)$.

Рассмотрим первый способ перехода в состояние $(1, 1)$. Вероятность состояния $(0, 1)$ на момент времени t обозначим $p_{01}(t)$. Вероятность того, что в систему поступит одно требование за время Δt , равна $\lambda \Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.10) с учетом вероятности $p_{01}(t)$ и переходной вероятности $\lambda \Delta t$:

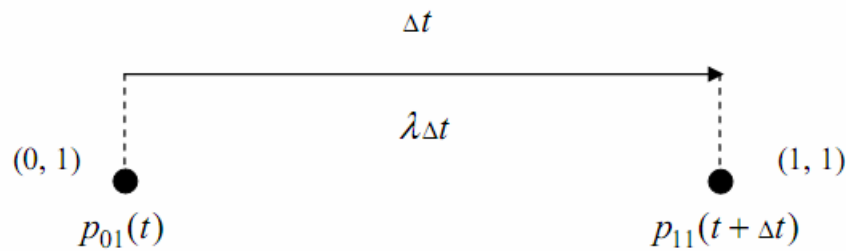


Рис. 5.10. Переход из состояния $(0, 1)$ в состояние $(1, 1)$

В соответствии с рис. 5.10 вероятность $p_{11}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{11}(t + \Delta t) = p_{01}(t) \lambda \Delta t. \quad (5.13)$$

Рассмотрим второй способ перехода в состояние $(1, 1)$. Вероятность состояния $(1, 1)$ на момент времени t обозначим $p_{11}(t)$. Вероятность того, что в системе за время Δt , не произойдет обслуживания, как в первой фазе, так и во второй, равна $1 - \mu \Delta t - \mu \Delta t = 1 - 2\mu \Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения второго

способа перехода (рис. 5.11) с учетом вероятности $p_{11}(t)$ и переходной вероятности $1 - 2\mu\Delta t$:

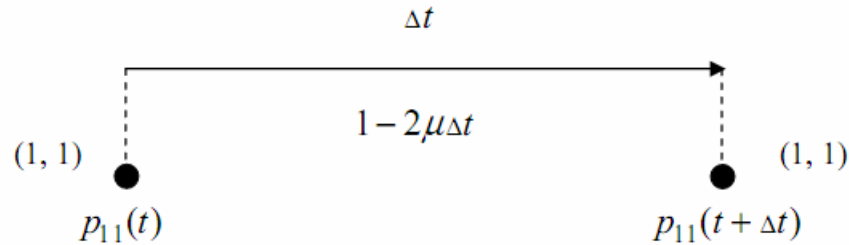


Рис. 5.11. Переход из состояния $(1, 1)$ в то же самое состояние $(1, 1)$

В соответствии с рис. 5.11 вероятность $p_{11}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{11}(t + \Delta t) = p_{11}(t)(1 - 2\mu\Delta t). \quad (5.14)$$

Для определения вероятности $p_{10}(t + \Delta t)$ с учетом (5.13), (5.14) применим формулу полной вероятности:

$$p_{11}(t + \Delta t) = p_{01}(t)\lambda\Delta t + p_{11}(t)(1 - 2\mu\Delta t). \quad (5.15)$$

Рассмотрим состояние $(b, 1)$ нахождения одного заблокированного требования в первой фазе и одного — во второй фазе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{b1}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если в системе, находящейся в состоянии $(1, 1)$, произойдет обслуживание в первой фазе: переход из состояния $(1, 1)$ в состояние $(b, 1)$,
- если в системе не произойдет обслуживания требования во второй фазе: формальный переход из состояния $(b, 1)$ в состояние $(b, 1)$.

Рассмотрим первый способ перехода в состояние $(b, 1)$. Вероятность состояния $(1, 1)$ на момент времени t обозначим $p_{11}(t)$. Вероятность того, что в системе за время Δt произойдет обслуживание требования в первой

фазе, равна $\mu\Delta t$, которую назовем переходной вероятностью. Графические пояснения первого способа перехода показаны на рис. 5.12:

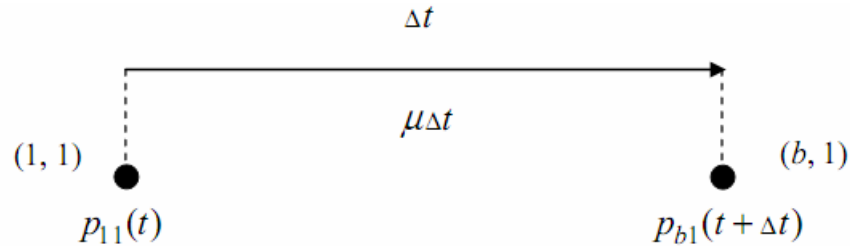


Рис. 5.12. Переход из состояния $(1, 1)$ в состояние $(b, 1)$

Вероятность $p_{b1}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{b1}(t + \Delta t) = p_{11}(t)\mu\Delta t. \quad (5.16)$$

Рассмотрим второй способ перехода в состояние $(b, 1)$. Вероятность состояния $(b, 1)$ на момент времени t обозначим $p_{b1}(t)$. Вероятность того, что в системе за время Δt не произойдет никаких изменений, т. е. не произойдет обслуживания требования во второй фазе, равна $1 - \mu\Delta t$, которую назовем переходной вероятностью. Пояснения второго способа перехода показаны на рис. 5.13.

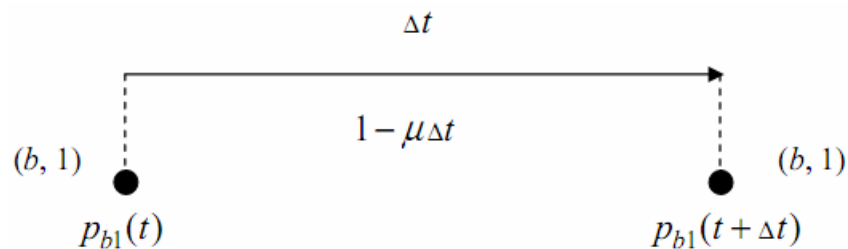


Рис. 5.13. Переход из состояния $(b, 1)$ в то же самое состояние $(b, 1)$

В соответствии с рис. 5.13 вероятность $p_{b1}(t + \Delta t)$ определим как произведение вероятностей:

$$p_{b1}(t + \Delta t) = p_{b1}(t)(1 - \mu\Delta t). \quad (5.17)$$

Для определения вероятности $p_{10}(t + \Delta t)$ с учетом (5.16), (5.17) применим формулу полной вероятности:

$$p_{b1}(t + \Delta t) = p_{11}(t)\mu\Delta t + p_{b1}(t)(1 - \mu\Delta t). \quad (5.18)$$

Таким образом, имеем следующие уравнения (5.4), (5.8), (5.12), (5.15), (5.18) первого приближения, в которых не учитываются малые более высокого порядка, чем Δt :

$$\left. \begin{aligned} p_{00}(t + \Delta t) &= p_{00}(t)(1 - \lambda\Delta t) + p_{01}(t)\mu\Delta t, \\ p_{01}(t + \Delta t) &= p_{01}(t)(1 - \mu\Delta t - \lambda\Delta t) + p_{10}(t)\mu\Delta t + p_{b1}(t)\mu\Delta t, \\ p_{10}(t + \Delta t) &= p_{00}(t)\lambda\Delta t + p_{10}(t)(1 - \mu\Delta t) + p_{11}(t)\mu\Delta t, \\ p_{11}(t + \Delta t) &= p_{01}(t)\lambda\Delta t + p_{11}(t)(1 - 2\mu\Delta t), \\ p_{b1}(t + \Delta t) &= p_{11}(t)\mu\Delta t + p_{b1}(t)(1 - \mu\Delta t). \end{aligned} \right\} \quad (5.19)$$

Преобразуем систему (5.19): раскроем скобки, приведем подобные члены, и перенесем в левую часть вероятности с теми нижними индексами, которыми обозначены вероятности левых частей, получим.

$$\left. \begin{aligned} p_{00}(t + \Delta t) - p_{00}(t) &= -p_{00}(t)\lambda\Delta t + p_{01}(t)\mu\Delta t, \\ p_{01}(t + \Delta t) - p_{01}(t) &= -(\mu\Delta t + \lambda\Delta t)p_{01}(t) + p_{10}(t)\mu\Delta t + p_{b1}(t)\mu\Delta t, \\ p_{10}(t + \Delta t) - p_{10}(t) &= p_{00}(t)\lambda\Delta t - p_{10}(t)\mu\Delta t + p_{11}(t)\mu\Delta t, \\ p_{11}(t + \Delta t) - p_{11}(t) &= p_{01}(t)\lambda\Delta t - 2p_{11}(t)\mu\Delta t, \\ p_{b1}(t + \Delta t) - p_{b1}(t) &= p_{11}(t)\mu\Delta t - p_{b1}(t)\mu\Delta t. \end{aligned} \right\} \quad (5.20)$$

Теперь поделим каждое из уравнений (5.20) на величину Δt и осуществим предельный переход в каждом их уравнений при $\Delta t \rightarrow 0$:

$$\left. \begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{p_{00}(t + \Delta t) - p_{00}(t)}{\Delta t} &= -\lim_{\Delta t \rightarrow 0} p_{00}(t)\lambda + \lim_{\Delta t \rightarrow 0} p_{01}(t)\mu, \\ \lim_{\Delta t \rightarrow 0} \frac{p_{01}(t + \Delta t) - p_{01}(t)}{\Delta t} &= -\lim_{\Delta t \rightarrow 0} (\mu + \lambda)p_{01}(t) + \lim_{\Delta t \rightarrow 0} p_{10}(t)\mu + \lim_{\Delta t \rightarrow 0} p_{b1}(t)\mu, \\ \lim_{\Delta t \rightarrow 0} \frac{p_{10}(t + \Delta t) - p_{10}(t)}{\Delta t} &= \lim_{\Delta t \rightarrow 0} p_{00}(t)\lambda - \lim_{\Delta t \rightarrow 0} p_{10}(t)\mu + \lim_{\Delta t \rightarrow 0} p_{11}(t)\mu, \\ \lim_{\Delta t \rightarrow 0} \frac{p_{11}(t + \Delta t) - p_{11}(t)}{\Delta t} &= \lim_{\Delta t \rightarrow 0} p_{01}(t)\lambda - 2 \lim_{\Delta t \rightarrow 0} p_{11}(t)\mu, \\ \lim_{\Delta t \rightarrow 0} \frac{p_{b1}(t + \Delta t) - p_{b1}(t)}{\Delta t} &= \lim_{\Delta t \rightarrow 0} p_{11}(t)\mu - \lim_{\Delta t \rightarrow 0} p_{b1}(t)\mu. \end{aligned} \right\} \quad (5.21)$$

Как можно видеть левая часть каждого из уравнений (5.21) формально представляет собой производную, в правой части предельные

соотношения существуют, т.к. они не зависят от Δt . В результате получаем следующую систему дифференциальных уравнений:

$$\begin{aligned}\frac{dp_{00}(t)}{dt} &= -\lambda p_{00}(t) + \mu p_{01}(t), \\ \frac{dp_{01}(t)}{dt} &= -(\mu + \lambda) p_{01}(t) + \mu p_{10}(t) + \mu p_{b1}(t), \\ \frac{dp_{10}(t)}{dt} &= \lambda p_{00}(t) - \mu p_{10}(t) + \mu p_{11}(t), \\ \frac{dp_{11}(t)}{dt} &= \lambda p_{01}(t) - 2\mu p_{11}(t), \\ \frac{dp_{b1}(t)}{dt} &= \mu p_{11}(t) - \mu p_{b1}(t).\end{aligned}\tag{5.22}$$

Полученная система (5.22) — это система обыкновенных дифференциальных уравнений с постоянными коэффициентами λ и μ . Для интегрирования системы (5.22) необходимо задать начальные условия. В качестве таких условий можно принять естественные начальные условия:

$$p_{00}(0) = 1, \quad p_{01}(0) = p_{10}(0) = p_{11}(0) = p_{b1}(0) = 0.\tag{5.23}$$

Условия (5.23) означают, что в начальный момент времени вероятность $p_{00}(0)$ отсутствия требований в системе равна единице, остальные вероятности возможных состояний равны нулю.

5.1.1. Стационарный режим двухфазной системы с нулевой вместимостью блока ожидания

Если в системе существует стационарный режим, то вероятности состояний в стационарном режиме можно получить из системы дифференциальных уравнений (5.22), приравняв к нулю производные в левых частях каждого из уравнений. В результате будем иметь

$$\begin{aligned}
 0 &= -\lambda p_{00} + \mu p_{01}, \\
 0 &= -(\mu + \lambda)p_{01} + \mu p_{10} + \mu p_{b1}, \\
 0 &= \lambda p_{00} - \mu p_{10} + \mu p_{11}, \\
 0 &= \lambda p_{01} - 2\mu p_{11}, \\
 0 &= \mu p_{11} - \mu p_{b1}.
 \end{aligned} \tag{5.24}$$

В алгебраической системе (5.24) входящие в неё вероятности состояний уже не зависят от времени. Преобразуем систему уравнений (5.24) так, чтобы выразить одни вероятности через другие:

$$\begin{aligned}
 0 &= -\lambda p_{00} + \mu p_{01}, \Rightarrow p_{01} = \rho p_{00}, \\
 0 &= -(\mu + \lambda)p_{01} + \mu p_{10} + \mu p_{b1}, \Rightarrow p_{10} = (1 + \rho)\rho p_{00} - p_{b1}, \\
 0 &= \lambda p_{00} - \mu p_{10} + \mu p_{11}, \Rightarrow p_{11} = -\rho p_{00} + (1 + \rho)\rho p_{00} - p_{b1}, \\
 0 &= \lambda p_{01} - 2\mu p_{11}, \Rightarrow p_{11} = 0.5\rho p_{01} = 0.5\rho^2 p_{00}, \\
 0 &= \mu p_{11} - \mu p_{b1}, \Rightarrow p_{b1} = p_{11},
 \end{aligned} \tag{5.25}$$

где $\rho = \lambda / \mu$. При этом одно из уравнений (3-е или 4-е) системы (5.25) является избыточным. Используя условие полной группы событий

$$p_{00} + p_{01} + p_{10} + p_{11} + p_{b1} = 1,$$

решение выше приведенной системы уравнений (5.25) будет иметь следующий вид:

$$p_{00} = \frac{1}{1.5\rho^2 + 2\rho + 1} = \frac{2}{3\rho^2 + 2\rho + 2}, \tag{5.26}$$

$$p_{01} = \frac{2\rho}{3\rho^2 + 2\rho + 2}, \tag{5.27}$$

$$p_{10} = \frac{\rho^2 + 2\rho}{3\rho^2 + 2\rho + 2}, \tag{5.28}$$

$$p_{11} = p_{b1} = \frac{\rho^2}{3\rho^2 + 2\rho + 2}. \tag{5.29}$$

5.1.2. Пример моделирования двухфазной системы с нулевой вместимостью блока ожидания

Примем следующие параметры системы: $\lambda = 12.5 \text{ с}^{-1}$, $\mu = 3.2 \text{ с}^{-1}$.

Зададим начальные условия для интегрирования системы дифференциальных уравнений (5.22):

$p_{00}(0) = 1$, $p_{01}(0) = p_{10}(0) = p_{11}(0) = p_{b1}(0) = 1$. Систему (5.22) представим в матричном виде

$$\frac{dP}{dt} = AP, \quad P(0) = [1, 0, 0, 0, 0]^T, \quad (5.30)$$

где $P = P(t)$ — вектор вероятностей состояний размера 5×1 , A — постоянная матрица коэффициентов размера 5×5 . $P(0)$ — вектор начальных условий.

Для заданных параметров матрица коэффициентов имеет вид

$$A = \begin{bmatrix} -12.5 & 3.2 & 0 & 0 & 0 \\ 0 & -15.7 & 3.2 & 0 & 3.2 \\ 12.5 & 0 & -3.2 & 3.2 & 0 \\ 0 & 12.5 & 0 & -6.4 & 0 \\ 0 & 0 & 0 & 3.2 & -3.2 \end{bmatrix}. \quad (5.31)$$

Для численного решения системы (5.30) с заданной матрицей (5.31) и визуализации результатов используем следующую MATLAB-программу:

```
function two;  
clc,format compact, format long  
close all  
L = 12.5;  
M = 3.2;  
  
A=[-L, M, 0, 0, 0;  
    0, -(L+M), M, 0, M;  
    L, 0, -M, M, 0;
```

```
0, L, 0, -2*M, 0;
0, 0, 0, M, -M];
P0 = [1;zeros(length(A)-1,1)];
T = [0, 4];
[t, P] = ode45(@test51,T,P0, [], A);
P_00 = P(end,1),P_01 = P(end,2),P_10 = P(end,3),
P_11 = P(end,4),P_b1 = P(end,5)
%%Control = P_00 + P_01 + P_10 + P_11 + P_b1
Control = sum(P(end,:))
plot(t,P,'linew',2.7),
grid on,
text(0.03,0.85,'\bf\fontsize{12}\leftarrow p_0_0(t)')
text(2.7,0.18,'\bf\fontsize{12} p_0_1(t)\newline\downarrow')
text(0.64,0.45,'\bf\fontsize{12}\leftarrow p_1_0(t)')
text(0.5,0.29,'\bf\fontsize{12}p_1_1(t)\newline\downarrow')
text(0.72,0.17,'\bf\fontsize{12}\leftarrow p_b_1(t)')
xlabel('\bf\it\fontsize{12} - - - - - t - - - - - '),
ylabel('\bf\fontsize{12} Вероятность'),
set(gcf,'color', [0.7,0.7,0.7]),

%% М-функция описания правых частей дифференциальных
уравнений:
function f = test51(t,P, A)
f = A*P;
```

Результат выполнения программы:

$P_{00} =$

0.031545059013963

$P_{01} =$

0.123223320161888

$P_{10} =$

0.363894139300362

$P_{11} =$

0.240669299427368

$P_{b1} =$

0.240668182096420

Control =

1.0000000000000000

Условие нормировки выполняется. Диаграмма вероятностей показана на рис. 5.14.

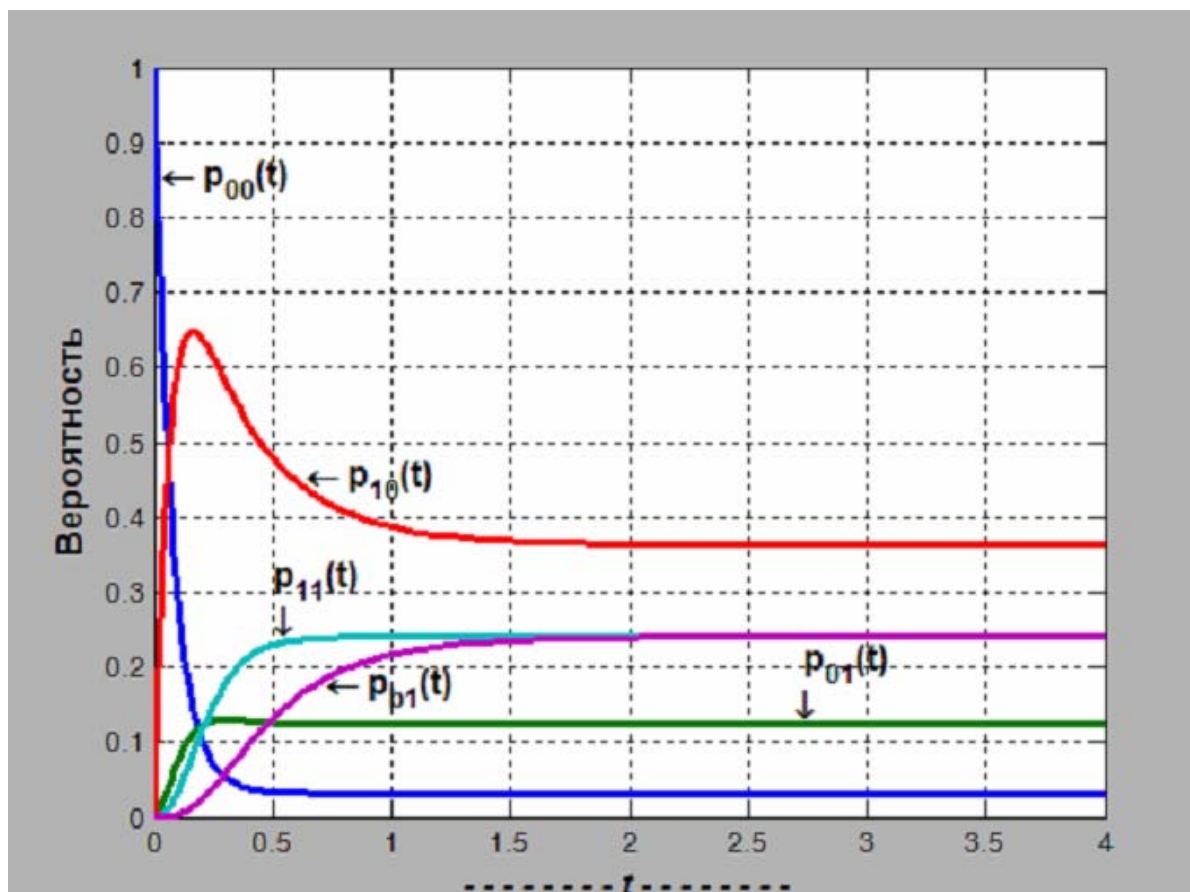


Рис. 5.14. Временные зависимости вероятностей двухфазной системы

Вычислим стационарные вероятности по формулам (5.26)-(5.29) и сравним со значениями установившихся вероятностей, полученными из решений системы дифференциальных уравнений:

$$\lambda = 12.5, \quad \mu = 3.2, \quad \rho = \lambda / \mu = 3.90625,$$

$$p_{00} = \frac{2}{3\rho^2 + 2\rho + 2} = 0.03154506107235,$$

$$p_{01} = \frac{2\rho}{3\rho^2 + 2\rho + 2} = 0.12322289481386,$$

$$p_{10} = \frac{\rho^2 + 2\rho}{3\rho^2 + 2\rho + 2} = 0.36389261124717,$$

$$p_{11} = p_{b1} = \frac{\rho^2}{3\rho^2 + 2\rho + 2} = 0.24066971643331.$$

Как видно, результаты совпадают с точностью до 5-6 знаков после запятой.

Определим еще некоторые операционные характеристики данной двухфазной системы:

1. Вероятность $P_{\phi 1}$ того, что вновь поступающее требование будет принято на обслуживание 1-й фазой:

$$P_{\phi 1} = p_{00} + p_{01} = 0.03154506107235 + 0.12322289481386 = 0.15476795588621.$$

2. Вероятность P_{omk} того, что вновь поступающее требование не будет принято на обслуживание 1-й фазой. Эту вероятность можно рассчитать как противоположную вероятности $P_{\phi 1}$:

$$P_{omk} = 1 - P_{\phi 1} = 0.52069176943547 \approx 0.52.$$

3. Относительная пропускная способность Q данной двухфазной системы:

$$Q = 1 - P_{omk} = 1 - 0.52 = 0.48.$$

С данными параметрами λ и μ система обладает невысокой пропускной способностью.

4. Эффективная частота $\lambda_{эфф}$ поступления требований в двухфазную систему:

$$\lambda_{эфф} = \lambda P_{\phi 1} = 1.9346 [c^{-1}].$$

5. Абсолютная пропускная способность A данной двухфазной системы:

$$A = \lambda_{эфф} Q = 1.9346 \cdot 0.48 = 0.928608 \approx 0.93 [c^{-1}].$$

Абсолютная пропускная способность показывает среднее число требований, обслуженных в единицу времени (здесь в одну секунду), данной двухфазной системой.

6. Среднее число требований N_s , находящихся в двухфазной системе.

Рассчитывается по формуле математического ожидания для дискретной случайной величины:

$$N_s = \sum_{k=0}^2 k P_k = 0 p_{00} + 1(p_{01} + p_{10}) + 2(p_{11} + p_{b1}) = 1.45,$$

где p_{00} — вероятность отсутствия требований в системе,

$p_{01} + p_{10}$ — вероятность того, что в системе находится одно требование,

$p_{11} + p_{b1}$ — вероятность того, что в системе находится два требования.

7. Полное время T_s пребывания требований (среднего числа) в двухфазной системе. Рассчитывается по формуле Литтла:

$$T_s = \frac{N_s}{\lambda_{эфф}} = 0.7494 [c].$$

8. Среднее время T_{cp} обработки одного требования в двухфазной системе:

$$T_{cp} = 1/\mu + 1/\mu = 2/\mu = 2/3.2 = 0.625 [c].$$

9. Средняя продолжительность задержки T_b требования в 1-й фазе, когда она заблокирована:

$$T_b = T_s - T_{cp} = 0.7494 - 0.625 = 0.1244 [c].$$

Рассмотренные операционные характеристики, как правило, являются достаточными для описания работоспособности двухфазной системы.

5.2. Трехфазная модель с нулевой вместимостью блоков ожидания

Схема рассматриваемой модели показана на рис. 5.15. Предполагается, что очередь перед каждой фазой не допускается.

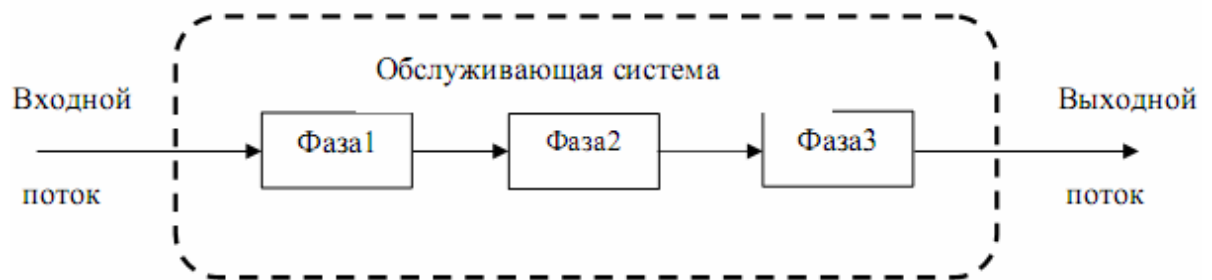


Рис. 5.15. Модель трехфазной системы обслуживания

Считаем, что входной поток требований пуассоновский со средней интенсивностью, равной λ . Обслуживание в каждой фазе осуществляется по экспоненциальному закону с одной и той же средней скоростью обслуживания μ . Каждое прибывающее требование должно сначала пройти через первую фазу обслуживания, затем через вторую и через третью. При этом принимается, что очереди перед каждой фазой не должны быть.

Рассмотрим возможные состояния системы, показанной на рис. 5.15. Для данной модели каждая из фаз может быть свободной (символ 0) и занятой (символ 1). При этом 1-я и 2-я фазы могут быть заблокированы (символ b). Блокировка 1-й фазы происходит в случае, если обслуживание в 1-й фазе завершено, а во 2-й фазе еще продолжается или 2-я фаза также заблокирована. Блокировка 2-й фазы происходит тогда, когда в ней обслуживание завершилось, а 3-я фаза еще находится в состоянии обслуживания. Состояние первой фазы будем обозначать символом i , состояние второй фазы — символом j , состояние третьей — символом k .

Возможны следующие состояния трехфазной системы обслуживания:

$$\{(i, j, k)\} = \left\{ \begin{array}{cccccc} (0,0,0) & (1,0,0) & (0,1,0) & (0,0,1) & (1,0,1) & (1,1,1) \\ (1,1,0) & (b,1,0) & (1,b,1) & (b,b,1) & (b,1,1) & (0,b,1) \end{array} \right\}. \quad (5.32)$$

Обозначим через $p_{ijk}(t)$ вероятность того, что в момент времени t система находится в состоянии (i, j, k) , т. е. в одном из возможных состояний в соответствии с (5.32).

Рассмотрим вероятности переходов из одного состояния в другое на временном интервале $[t, t + \Delta t]$, где Δt — бесконечно малое положительное приращение времени. Для удобства создадим таблицу 5.2, в которой левый столбец будет представлять исходные состояния системы в момент времени t , а в верхней строке будут указаны состояния (состояния на момент времени $[t + \Delta t]$), в которые за время Δt система может перейти или не перейти. Вероятности возможных переходов за время Δt — переходные вероятности — указаны в остальных ячейках таблицы. Число возможных состояний системы равно 13. Число переходных вероятностей равно 36. Таблица имеет 169 ячеек, 36 из которых заполнены возможными переходными вероятностями. Если

переход из одного состояния в другое невозможен, то в соответствующей ячейке таблицы ставится прочерк. Столбец таблицы с заголовком "Состояния в момент t " представляет собой возможные исходные состояния системы — 13 состояний. В строке таблицы с заголовком "Состояния системы на момент времени $[t+\Delta t]$ " указаны возможные состояния системы на момент времени $[t+\Delta t]$ — 13 состояний. Переходные вероятности, расположенные в 36 ячейках таблицы, это вероятности, которые осуществляют перевод системы из одного состояния в другое за малый отрезок времени Δt .

Таблица 5.2

Состояния системы на момент времени $[t+\Delta t]$											
(000)	(100)	(010)	(001)	(101)	(011)	(111)	(110)	(b10)	(1b1)	(bb1)	(0b1)
Переходные вероятности за время Δt											
$1-\lambda\Delta t$	$\lambda\Delta t$	—	—	—	—	—	—	—	—	—	—
—	$1-\mu\Delta t$	$(1-\lambda\Delta t)*\mu\Delta t$	—	—	—	—	—	—	—	—	—
—	—	$(1-\lambda\Delta t)*\mu\Delta t$	$(1-\lambda\Delta t)*\mu\Delta t$	—	—	—	$\lambda\Delta t(1-\mu\Delta t)$	—	—	—	—
$(1-\lambda\Delta t)*\mu\Delta t$	—	—	$(1-\lambda\Delta t)*\mu\Delta t$	$\lambda\Delta t(1-\mu\Delta t)$	—	—	—	—	—	—	—
—	$(1-\mu\Delta t)\mu\Delta t$	—	—	$(1-\mu\Delta t)^2$	$(1-\lambda\Delta t)*\mu\Delta t(1-\mu\Delta t)$	—	—	—	—	—	—
—	—	$(1-\lambda\Delta t)*\mu\Delta t$	—	—	$(1-\lambda\Delta t)*\mu\Delta t(1-\mu\Delta t)^2$	$\lambda\Delta t*$	—	—	—	—	$(1-\lambda\Delta t)*\mu\Delta t*$
—	—	—	—	—	—	$(1-\mu\Delta t)^3$	$\mu\Delta t$	—	$(1-\mu\Delta t)\mu\Delta t*$	$\mu\Delta t(1-\mu\Delta t)$	—
—	—	—	—	$\mu\Delta t$	—	—	$(1-\mu\Delta t)^2$	$\mu\Delta t(1-\mu\Delta t)$	—	—	—
—	—	—	—	—	$(1-\lambda\Delta t)\mu\Delta t$	—	—	$(1-\mu\Delta t)$	—	—	—
—	—	—	—	$\mu\Delta t$	—	—	—	—	$(1-\mu\Delta t)^2$	$\mu\Delta t(1-\mu\Delta t)$	—
—	—	—	—	—	$(1-\lambda\Delta t)\mu\Delta t$	—	—	—	—	$(1-\mu\Delta t)$	—
—	—	—	—	—	—	—	—	$(1-\mu\Delta t)\mu\Delta t$	—	$\mu\Delta t(1-\mu\Delta t)$	—
—	—	—	$(1-\lambda\Delta t)*\mu\Delta t$	—	—	—	—	—	$\lambda\Delta t(1-\mu\Delta t)$	—	$(1-\lambda\Delta t)*\mu\Delta t$

Состоя- ния в момент t	(000)	(100)	(010)	(001)	(101)	(011)	(111)	(110)	(b10)	(1b1)	(bb1)	(b11)	(0b1)
-----------------------------------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Определим вероятности всех возможных состояний на момент времени $[t + \Delta t]$, пренебрегая малыми более высокого порядка чем Δt .

1. Рассмотрим состояние (0,0,0) отсутствия требований в системе на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{000}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если в систему не поступит ни одного требования: переход из состояния (0,0,0) в состояние (0,0,0);
- если в системе было одно требование в третьей фазе, это требование было обслужено, и оно покинуло систему, но в систему не поступит больше требования: переход из состояния (0,0,1) в состояние (0,0,0).

Рассмотрим первый способ перехода в состояние (0,0,0). Вероятность отсутствия требований на момент времени t обозначим $p_{000}(t)$. Вероятность того, что требование не поступило за время Δt определяется по формуле (5.2) как $(1 - \lambda \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.16) с учетом вероятности $p_{000}(t)$ и переходной вероятности $(1 - \lambda \Delta t)$:

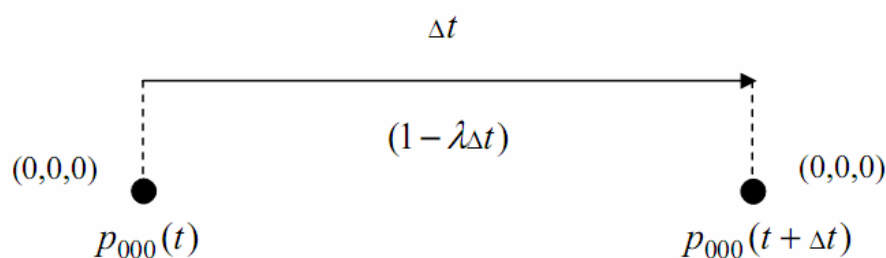


Рис. 5.16. Переход из состояния (0,0,0) в состояние (0,0,0) за время Δt

Учитывая, что пуассоновский поток обладает свойством отсутствия последствия, запишем на основе теоремы умножения вероятностей для рассматриваемого перехода следующее соотношение (см. рис. 5.16):

$$p_{000}(t + \Delta t) = p_{000}(t)(1 - \lambda \Delta t). \quad (5.33)$$

Рассмотрим второй способ перехода в состояние $(0,0,0)$. Для этого также выполним графические пояснения второго способа перехода (рис. 5.17) с учетом вероятности $p_{001}(t)$ и переходной вероятности $(1 - \lambda \Delta t)\mu \Delta t$:

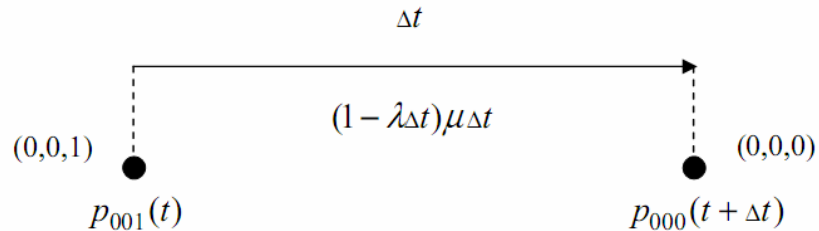


Рис. 5.17. Переход из состояния $(0,0,1)$ в состояние $(0,0,0)$ за время Δt

В соответствии с рис. 5.17 запишем соотношение, связывающее вероятность $p_{000}(t + \Delta t)$ с переходной вероятностью $(1 - \lambda \Delta t)\mu \Delta t$ и вероятностью $p_{001}(t)$:

$$p_{000}(t + \Delta t) = p_{001}(t)(1 - \lambda \Delta t)\mu \Delta t. \quad (5.34)$$

Рассмотренные переходы в состояние $(0,0,0)$ за время Δt являются событиями несовместными, поэтому применим формулу полной вероятности, учитывая (5.33) и (5.34):

$$p_{000}(t + \Delta t) = p_{000}(t)(1 - \lambda \Delta t) + p_{001}(t)(1 - \lambda \Delta t)\mu \Delta t. \quad (5.35)$$

2. Рассмотрим состояние $(1,0,0)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{100}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если в систему поступит одно требование: переход из состояния $(0,0,0)$ в состояние $(1,0,0)$;
- если система не изменила своего состояния: переход из состояния $(1,0,0)$ в состояние $(1,0,0)$;
- если первая фаза не обслужила требования, а из третьей фазы вышло обслуженное требование: переход из состояния $(1,0,1)$ в состояние $(1, 0,0)$.

Рассмотрим первый способ перехода в состояние $(1,0,0)$. Вероятность того, что в систему поступит одно требование за время Δt определяется как $\lambda \Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.18) с учетом вероятности $p_{000}(t)$ и переходной вероятности $\lambda \Delta t$:

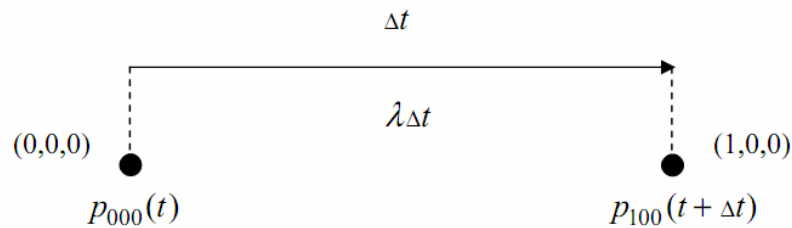


Рис. 5.18. Переход из состояния $(0,0,0)$ в состояние $(1,0,0)$ за время Δt

Запишем соотношение, связывающее вероятность $p_{100}(t + \Delta t)$ с переходной вероятностью $\lambda \Delta t$ и вероятностью $p_{000}(t)$ в соответствии с теоремой умножения вероятностей:

$$p_{100}(t + \Delta t) = p_{000}(t) \lambda \Delta t. \quad (5.36)$$

Рассмотрим второй способ перехода в состояние $(1,0,0)$. Для этого выполним графические пояснения второго способа перехода (рис. 5.19) с учетом вероятности $p_{100}(t)$ и переходной вероятности $1 - \mu \Delta t$:

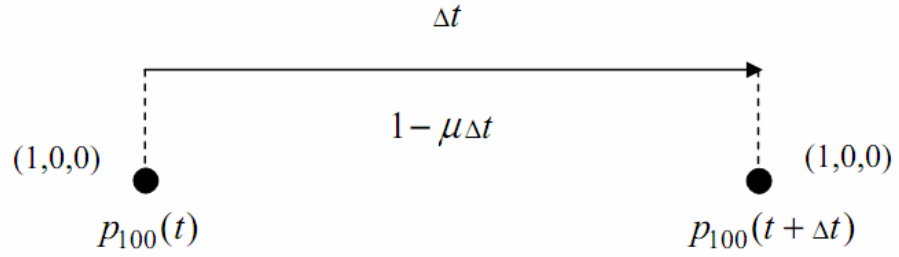


Рис. 5.19. Переход из состояния (1,0,0) в состояние (1,0,0) за время Δt

Запишем соотношение, связывающее вероятность $p_{100}(t + \Delta t)$ с переходной вероятностью $1 - \mu\Delta t$ и вероятностью $p_{100}(t)$:

$$p_{100}(t + \Delta t) = p_{100}(t)(1 - \mu\Delta t). \quad (5.37)$$

Рассмотрим третий способ перехода в состояние (1,0,0). Для этого выполним графические пояснения третьего способа перехода (рис. 5.20) с учетом вероятности $p_{101}(t)$ и переходной вероятности $(1 - \mu\Delta t)\mu\Delta t$:

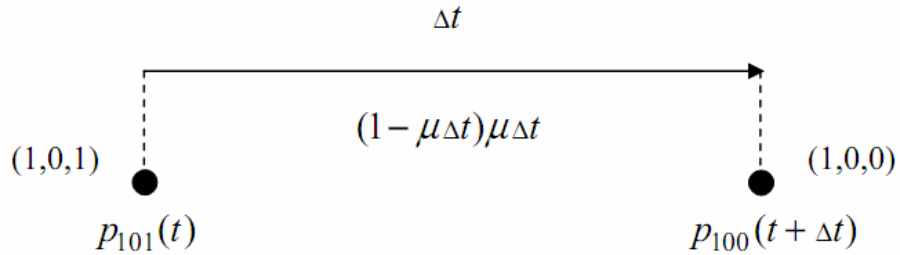


Рис. 5.20. Переход из состояния (1,0,1) в состояние (1,0,0) за время Δt

Запишем соотношение, связывающее вероятность $p_{100}(t + \Delta t)$ с переходной вероятностью $(1 - \mu\Delta t)\mu\Delta t$ и вероятностью $p_{101}(t)$:

$$p_{100}(t + \Delta t) = p_{101}(t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.38)$$

Рассмотренные три способа перехода в состояние (1,0,0) являются несовместными, поэтому можно применить формулу полной вероятности для несовместных событий с учетом (5.36), (5.37), (5.38):

$$p_{100}(t + \Delta t) = p_{000}(t)\lambda\Delta t + p_{100}(t)(1 - \mu\Delta t) + p_{101}(t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.39)$$

3. Рассмотрим состояние (0,1,0) системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{010}(t + \Delta t)$. От

момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если в первой фазе будет обслужено требование, при условии, что в систему не поступит требования: переход из состояния $(1,0,0)$ в состояние $(0,1,0)$;
- если в систему не поступит требования, а во второй фазе не произойдет обслуживания: переход из состояния $(0,1,0)$ в то же самое состояние $(0,1,0)$.
- если в систему не поступит требование (в 1 фазу), не произойдет обслуживания во 2-й фазе, а 3-я фаза освободится, в ней завершится обслуживание требования.

Рассмотрим первый способ перехода в состояние $(0,1,0)$. Вероятность того, что за время Δt в первой фазе произойдет обслуживание требования определяется как $(1 - \lambda\Delta t)\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.21) с учетом вероятности $p_{100}(t)$ и переходной вероятности $(1 - \lambda\Delta t)\mu\Delta t$:

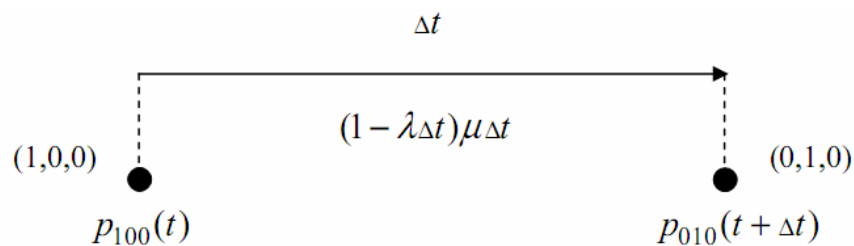


Рис. 5.21. Переход из состояния $(1,0,0)$ в состояние $(0,1,0)$ за время Δt
При переходе системы из состояния $(1,0,0)$ в состояние $(0,1,0)$ расчет вероятности $p_{010}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{010}(t + \Delta t) = p_{100}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.40)$$

Рассмотрим второй способ перехода в состояние $(0,1,0)$. Вероятность того, что за время Δt в систему не поступит требования и не произойдет обслуживания во второй фазе, определяется как $(1 - \lambda \Delta t)(1 - \mu \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.22) с учетом вероятности $p_{010}(t)$ и переходной вероятности $(1 - \lambda \Delta t)(1 - \mu \Delta t)$:

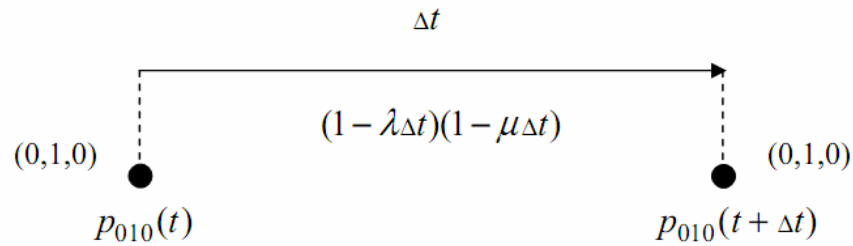


Рис. 5.22. Переход из состояния $(0,1,0)$ в состояние $(0,1,0)$ за время Δt

При переходе системы из состояния $(0,1,0)$ в состояние $(0,1,0)$ расчет вероятности $p_{010}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{010}(t + \Delta t) = p_{010}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t). \quad (5.41)$$

Рассмотрим третий способ перехода в состояние $(0,1,0)$. Вероятность того, что за время Δt в систему не поступит требования, не произойдет обслуживания во второй фазе и освободится третья фаза, определяется как $(1 - \lambda \Delta t)(1 - \mu \Delta t)\mu \Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.23) с учетом вероятности $p_{010}(t)$ и переходной вероятности $(1 - \lambda \Delta t)(1 - \mu \Delta t)$:

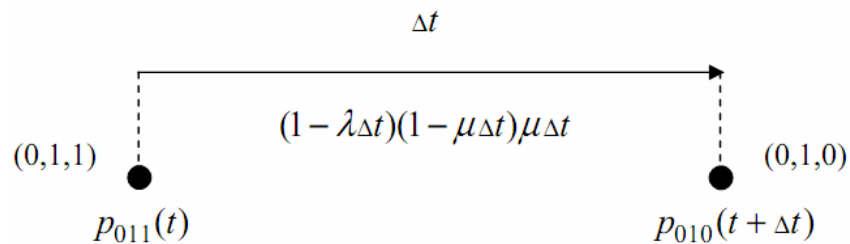


Рис. 5.23. Переход из состояния $(0,1,1)$ в состояние $(0,1,0)$ за время Δt

При переходе системы из состояния $(0,1,1)$ в состояние $(0,1,0)$ расчет вероятности $p_{010}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{010}(t + \Delta t) = p_{011}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.42)$$

Три рассмотренных способа перехода системы в состояние $(0,1,0)$ являются событиями несовместными, поэтому для определения вероятности $p_{010}(t + \Delta t)$ с учетом трех способов перехода (5.40), (5.41), (5.42) применим формулу полной вероятности для несовместных событий:

$$p_{010}(t + \Delta t) = p_{100}(t)(1 - \lambda\Delta t)\mu\Delta t + p_{010}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t) + p_{011}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.43)$$

4. Рассмотрим состояние $(0,0,1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{001}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если за время Δt вторая фаза обслужит требование и в систему не поступит новое требование: переход из состояния $(0,1,0)$ в состояние $(0,0,1)$;
- если за время Δt третья фаза не обслужит требования и в систему не поступит требования: переход из состояния $(0,0,1)$ в то же самое состояние $(0,0,1)$;
- если за время Δt третья фаза обслужит требование и в систему не поступит новое требование: переход из состояния $(0,b,1)$ в состояние $(0,0,1)$.

Рассмотрим первый способ перехода в состояние $(0,0,1)$. Вероятность того, что за время Δt второй фазе произойдет обслуживание требования, определяется как $(1 - \lambda\Delta t)\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа

перехода (рис. 5.24) с учетом вероятности $p_{010}(t)$ и переходной вероятности $(1 - \lambda\Delta t)\mu\Delta t$:

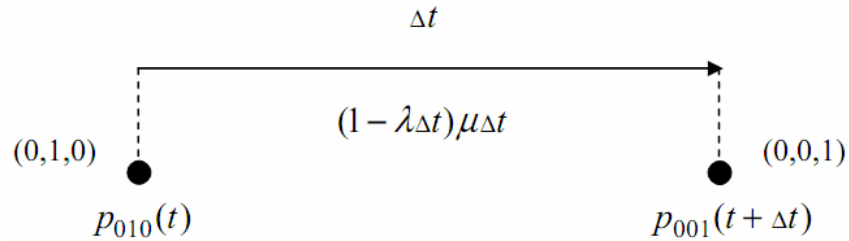


Рис. 5.24. Переход из состояния $(0,1,0)$ в состояние $(0,0,1)$ за время Δt

При переходе системы из состояния $(0,1,0)$ в состояние $(0,0,1)$ расчет вероятности $p_{001}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{001}(t + \Delta t) = p_{010}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.44)$$

Рассмотрим второй способ перехода в состояние $(0,0,1)$. Вероятность того, что за время Δt не произойдет обслуживания требования в третьей фазе, определяется как $(1 - \lambda\Delta t)(1 - \mu\Delta t)$. Выполним графические пояснения второго способа перехода (рис. 5.25) с учетом $p_{001}(t)$ и переходной вероятности $(1 - \lambda\Delta t)(1 - \mu\Delta t)$:

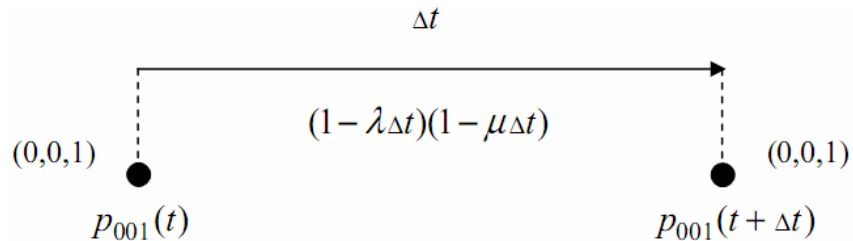


Рис. 5.25. Переход из состояния $(0,0,1)$ в состояние $(0,0,1)$ за время Δt

При переходе системы из состояния $(0,0,1)$ в то же самое состояние $(0,0,1)$ расчет вероятности $p_{001}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{001}(t + \Delta t) = p_{001}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t). \quad (5.45)$$

Рассмотрим третий способ перехода в состояние $(0,0,1)$. Вероятность того, что за время Δt произойдет обслуживание требования в третьей

фазе и не поступит нового требования в систему, равно $(1 - \lambda\Delta t)\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.26) с учетом $p_{0b1}(t)$ и переходной вероятности $(1 - \lambda\Delta t)\mu\Delta t$:

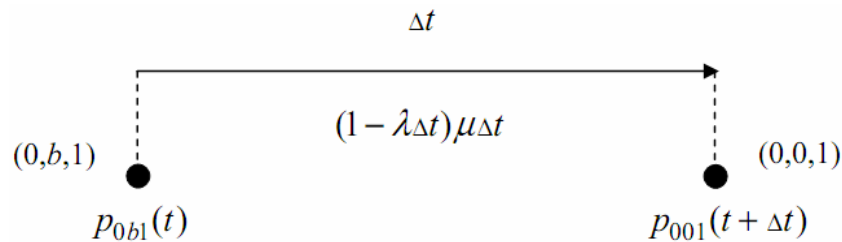


Рис. 5.26. Переход из состояния $(0, b, 1)$ в состояние $(0, 0, 1)$ за время Δt

При переходе системы из состояния $(0, b, 1)$ в состояние $(0, 0, 1)$ расчет вероятности $p_{001}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{001}(t + \Delta t) = p_{0b1}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.46)$$

Все три рассмотренные способы перехода в состояние $(0, 0, 1)$ являются несовместными, поэтому для определения вероятности $p_{001}(t + \Delta t)$ применим формулу полной вероятности с учетом (5.44), (5.45), (5.46):

$$p_{001}(t + \Delta t) = p_{010}(t)(1 - \lambda\Delta t)\mu\Delta t + p_{001}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t) + p_{0b1}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.47)$$

5. Рассмотрим состояние $(1, 0, 1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{101}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти четырьмя способами:

- если в систему поступит требование (в первую фазу) и если не произойдет обслуживания требования в третьей фазе: переход системы из состояния $(0, 0, 1)$ в состояние $(1, 0, 1)$;
- если в системе, находящейся в состоянии $(1, 0, 1)$, не произойдет изменений, т. е. не произойдет обслуживания в первой и третьей

фазах: переход системы из состояния $(1,0,1)$ в то же самое состояние $(1,0,1)$;

- если во второй фазе произойдет обслуживание требования: переход системы из состояния $(1,1,0)$ в состояние $(1,0,1)$;
- если в третьей фазе произойдет обслуживание требования: переход системы из состояния $(1,b,1)$ в состояние $(1,0,1)$.

Рассмотрим первый способ перехода в состояние $(1,0,1)$. Вероятность того, что за время Δt в систему поступит требование и не произойдет обслуживания требования в третьей фазе, определяется как $\lambda\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.27) с учетом вероятности $p_{001}(t)$ и переходной вероятности $\lambda\Delta t(1 - \mu\Delta t)$:

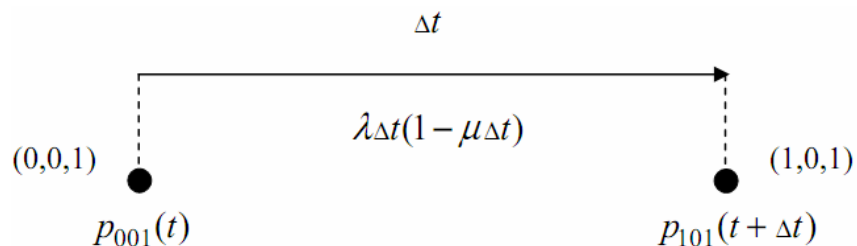


Рис. 5.27. Переход из состояния $(0,0,1)$ в состояние $(1,0,1)$ за время Δt

При переходе системы из состояния $(0,0,1)$ в состояние $(1,0,1)$ расчет вероятности $p_{101}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{101}(t + \Delta t) = p_{001}(t)\lambda\Delta t(1 - \mu\Delta t). \quad (5.48)$$

Рассмотрим второй способ перехода в состояние $(1,0,1)$. Вероятность того, что за время Δt не произойдет обслуживания ни в первой фазе, ни в третьей, определяется как $(1 - \mu\Delta t)(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.28) с учетом вероятности $p_{101}(t)$ и переходной вероятности $(1 - \mu\Delta t)(1 - \mu\Delta t)$:

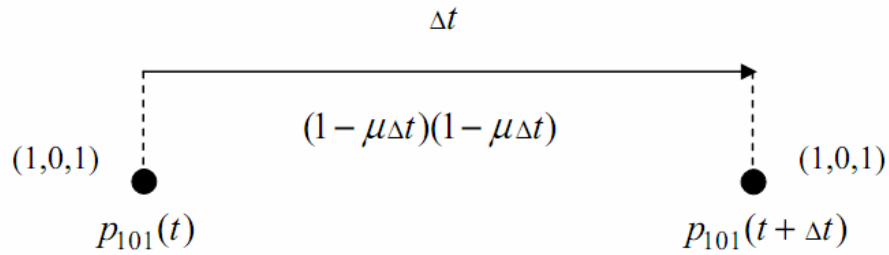


Рис. 5.28. Переход из состояния (1,0,1) в состояние (1,0,1) за время Δt

При переходе системы из состояния (1,0,1) в состояние (1,0,1) расчет вероятности $p_{101}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{101}(t + \Delta t) = p_{101}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.49)$$

Рассмотрим третий способ перехода в состояние (1,0,1). Вероятность того, что за время Δt произойдет обслуживание во второй фазе, определяется как $\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.29) с учетом вероятности $p_{110}(t)$ и переходной вероятности $\mu\Delta t$:

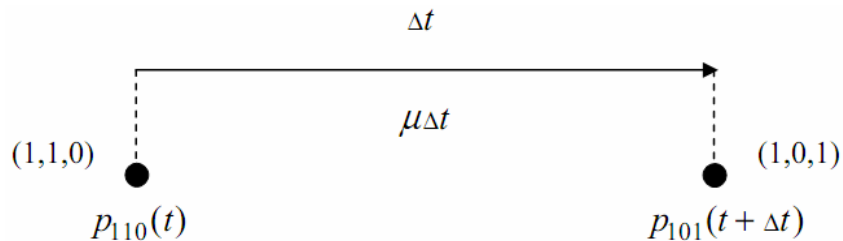


Рис. 5.29. Переход из состояния (1,1,0) в состояние (1,0,1) за время Δt

При переходе системы из состояния (1,1,0) в состояние (1,0,1) расчет вероятности $p_{101}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{101}(t + \Delta t) = p_{110}(t)\mu\Delta t. \quad (5.50)$$

Рассмотрим четвертый способ перехода в состояние (1,0,1). Вероятность того, что за время Δt произойдет обслуживание в третьей фазе, определяется как $\mu\Delta t$, которую назовем переходной вероятностью.

Выполним графические пояснения четвертого способа перехода (рис. 5.30) с учетом вероятности $p_{1b1}(t)$ и переходной вероятности $\mu\Delta t$:

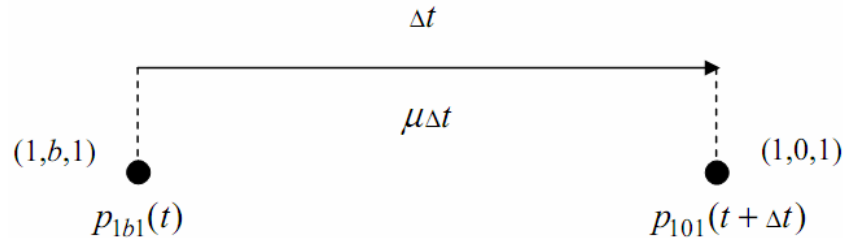


Рис. 5.30. Переход из состояния $(1, \mathbf{b}, 1)$ в состояние $(1, 0, 1)$ за время Δt

При переходе системы из состояния $(1, \mathbf{b}, 1)$ в состояние $(1, 0, 1)$ расчет вероятности $p_{101}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{101}(t + \Delta t) = p_{1b1}(t)\mu\Delta t. \quad (5.51)$$

Рассмотренные четыре способа перехода в состояние $(1, 0, 1)$ являются несовместными событиями, поэтому для расчета вероятности $p_{101}(t + \Delta t)$ с учетом всех возможных переходов (5.48), (5.49), (5.50), (5.51) можно использовать формулу полной вероятности:

$$p_{101}(t + \Delta t) = p_{001}(t)\lambda\Delta t(1 - \mu\Delta t) + p_{101}(t)(1 - \mu\Delta t)(1 - \mu\Delta t) + p_{110}(t)\mu\Delta t + p_{1b1}(t)\mu\Delta t. \quad (5.52)$$

6. Рассмотрим состояние $(0, 1, 1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{011}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти четырьмя способами:

- если в 1-й фазе произойдет обслуживание требования, в 3-й фазе обслуживания требования не произойдет и в систему не поступит нового требования: переход системы из состояния $(1, 0, 1)$ в состояние $(0, 1, 1)$;
- если система не изменит своего состояния, т.е. если в 1-ю фазу не поступит требования, а 2-я и 3-я фазы будут заняты обслуживанием

- требований: переход системы из состояния $(0,1,1)$ в то же самое состояние $(0,1,1)$;
- если 2-я фаза осуществит обслуживание требования и в систему не поступит нового требования: переход системы из состояния $(b,1,0)$ в состояние $(0,1,1)$;
- если 3-я фаза осуществит обслуживание требования и в систему не поступит нового требования: переход системы из состояния $(b,b,1)$ в состояние $(0,1,1)$.

Рассмотрим первый способ перехода в состояние $(0,1,1)$. Вероятность того, что за время Δt произойдет обслуживание в 1-й фазе, не произойдет обслуживания требования в 3-й фазе и не поступит в систему новое требование, определяется как $(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.31) с учетом вероятности $p_{101}(t)$ и переходной вероятности $(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t)$:

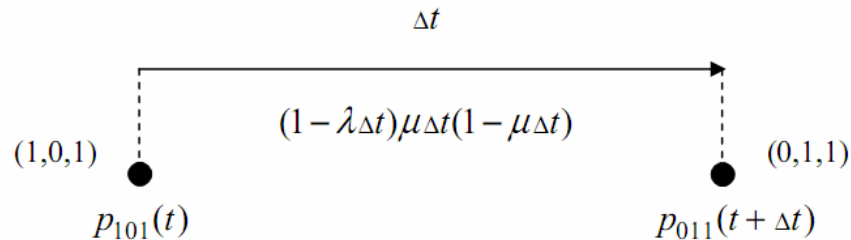


Рис. 5.31. Переход из состояния $(1,0,1)$ в состояние $(0,1,1)$ за время Δt

При переходе системы из состояния $(1,0,1)$ в состояние $(0,1,1)$ расчет вероятности $p_{011}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{011}(t + \Delta t) = p_{101}(t)(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t). \quad (5.53)$$

Рассмотрим второй способ перехода в состояние $(0,1,1)$. Вероятность того, что за время Δt в систему (в 1-ю фазу) не поступит требования, а 2-я и 3-я фазы будут заняты обслуживанием требований, т.е. в них не

произойдет обслуживания, определяется как $(1 - \lambda\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t)$. Выполним графические пояснения второго способа перехода (рис. 5.32) с учетом вероятности $p_{011}(t)$ и переходной вероятности $(1 - \lambda\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t)$:

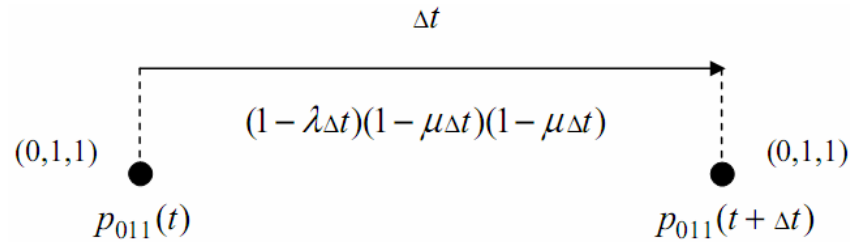


Рис. 5.32. Переход из состояния $(0,1,1)$ в состояние $(0,1,1)$ за время Δt

При переходе системы из состояния $(0,1,1)$ в то же самое состояние расчет $p_{011}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{011}(t + \Delta t) = p_{011}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.54)$$

Рассмотрим третий способ перехода в состояние $(0,1,1)$. Вероятность того, что за время Δt вторая фаза обслужит требование и в систему не поступит новое требование, определяется как $(1 - \lambda\Delta t)\mu\Delta t$. Выполним графические пояснения третьего способа перехода (рис. 5.33) с учетом вероятности $p_{b10}(t)$ и переходной вероятности $(1 - \lambda\Delta t)\mu\Delta t$:

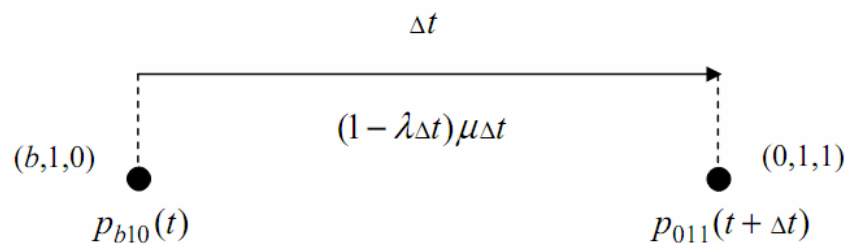


Рис. 5.33. Переход из состояния $(b,1,0)$ в состояние $(0,1,1)$ за время Δt

При переходе системы из состояния $(b,1,0)$ в состояние $(0,1,1)$ расчет вероятности $p_{011}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{011}(t + \Delta t) = p_{b10}(1 - \lambda\Delta t)\mu\Delta t. \quad (5.55)$$

Рассмотрим четвертый способ перехода в состояние $(0,1,1)$. Вероятность того, что за время Δt третья фаза обслужит требование, определяется как $(1 - \lambda\Delta t)\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения четвертого способа перехода (рис. 5.34) с учетом вероятности $p_{bb1}(t)$ и переходной вероятности $(1 - \lambda\Delta t)\mu\Delta t$:

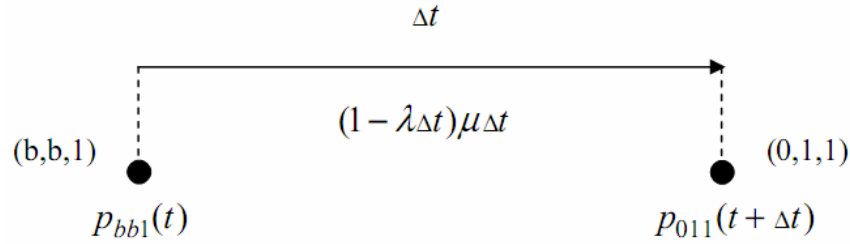


Рис. 5.34. Переход из состояния $(b,b,1)$ в состояние $(0,1,1)$ за время Δt

При переходе системы из состояния $(b,b,1)$ в состояние $(0,1,1)$ расчет вероятности $p_{011}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{011}(t + \Delta t) = p_{bb1}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.56)$$

Рассмотренные четыре способа перехода в состояние $(0,1,1)$ являются несовместными событиями, поэтому для расчета вероятности $p_{011}(t + \Delta t)$ с учетом всех возможных переходов (5.53), (5.54), (5.55), (5.56) можно использовать формулу полной вероятности:

$$p_{011}(t + \Delta t) = p_{101}(t)(1 - \lambda\Delta t)\mu\Delta t(1 - \mu\Delta t) + p_{011}(t)(1 - \lambda\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t) + p_{b10}(1 - \lambda\Delta t)\mu\Delta t + p_{bb1}(t)(1 - \lambda\Delta t)\mu\Delta t. \quad (5.57)$$

7. Рассмотрим состояние $(1,1,1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим $p_{111}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если в систему не поступит требования, а во второй и третьей фазах не произойдет обслуживание требований: переход системы из состояния $(0,1,1)$ в состояние $(1,1,1)$;
- если не произойдет обслуживание требований ни в первой, ни во второй, ни в третьей фазах: переход системы из состояния $(1,1,1)$ в то же самое состояние $(1,1,1)$.

Рассмотрим первый способ перехода в состояние $(1,1,1)$. Вероятность того, что за время Δt в систему поступит требование, а во второй и третьей фазах не произойдет обслуживания требований, определяется как $\lambda \Delta t (1 - \mu \Delta t)(1 - \mu \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.35) с учетом вероятности $p_{011}(t)$ и переходной вероятности $\lambda \Delta t (1 - \mu \Delta t)(1 - \mu \Delta t)$:

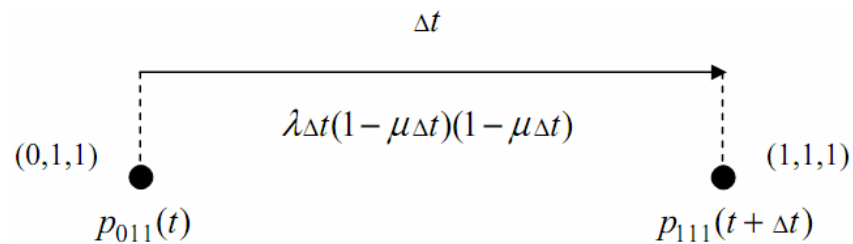


Рис. 5.35. Переход из состояния $(0,1,1)$ в состояние $(1,1,1)$ за время Δt

При переходе системы из состояния $(0,1,1)$ в состояние $(1,1,1)$ расчет вероятности $p_{111}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{111}(t + \Delta t) = p_{011}(t) \lambda \Delta t (1 - \mu \Delta t)(1 - \mu \Delta t). \quad (5.58)$$

Рассмотрим второй способ перехода в состояние $(1,1,1)$. Вероятность того, что за время Δt не произойдет изменений в системе, определяется как $(1 - \mu \Delta t)(1 - \mu \Delta t)(1 - \mu \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения рассматриваемого перехода (рис. 5.36)

с учетом вероятности $p_{111}(t)$ и переходной вероятности $(1 - \mu\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t)$:

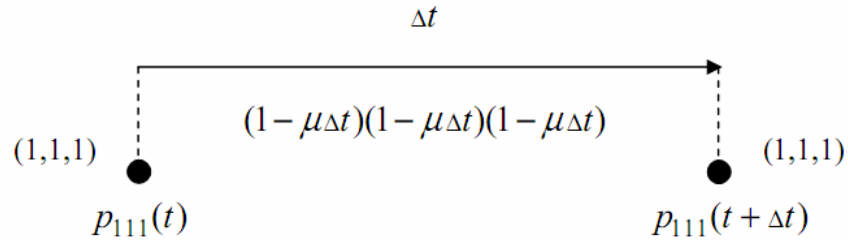


Рис. 5.36. Переход из состояния $(1,1,1)$ в состояние $(1,1,1)$ за время Δt

При переходе системы из состояния $(1,1,1)$ в то же самое состояние $(1,1,1)$ расчет вероятности $p_{111}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{111}(t + \Delta t) = p_{111}(t)(1 - \mu\Delta t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.59)$$

Рассмотренные два способа перехода в состояние $(1,1,1)$ являются несовместными событиями, поэтому для расчета вероятности $p_{111}(t + \Delta t)$ с учетом всех возможных переходов (5.58), (5.59) можно использовать формулу полной вероятности:

$$p_{111}(t + \Delta t) = p_{011}(t)\lambda\Delta t(1 - \mu\Delta t)(1 - \mu\Delta t) + p_{111}(t)(1 - \mu\Delta t)^3. \quad (5.60)$$

8. Рассмотрим состояние $(1,1,0)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{110}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если в систему поступит одно требование в первую фазу и не произойдет обслуживания во второй фазе: переход системы из состояния $(0,1,0)$ в состояние $(1,1,0)$;
- если произойдет обслуживание в третьей фазе: переход системы из состояния $(1,1,1)$ в состояние $(1,1,0)$;

— если не произойдет обслуживания требований ни в первой фазе, ни во второй: переход системы из состояния $(1,1,0)$ в то же самое состояние $(1,1,0)$.

Рассмотрим первый способ перехода в состояние $(1,1,0)$. Вероятность того, что за время Δt в систему поступит одно требование и не произойдет обслуживания во второй фазе, определяется как $\lambda\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.37) с учетом вероятности $p_{010}(t)$ и переходной вероятности $\lambda\Delta t(1 - \mu\Delta t)$:

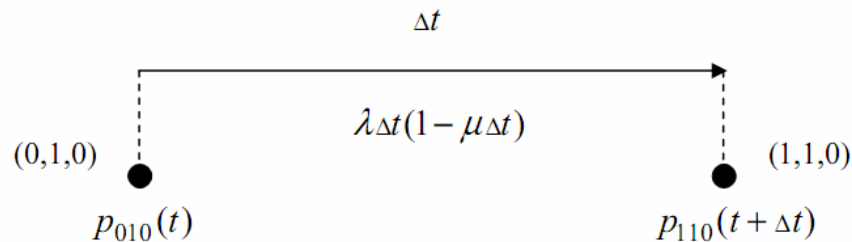


Рис. 5.37. Переход из состояния $(0,1,0)$ в состояние $(1,1,0)$ за время Δt

При переходе системы из состояния $(0,1,0)$ в состояние $(1,1,0)$ расчет вероятности $p_{110}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{110}(t + \Delta t) = p_{010}(t)\lambda\Delta t(1 - \mu\Delta t). \quad (5.61)$$

Рассмотрим второй способ перехода в состояние $(1,1,0)$. Вероятность того, что за время Δt произойдет обслуживание требования в третьей фазе, определяется как $\mu\Delta t$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.38) с учетом вероятности $p_{111}(t)$ и переходной вероятности $\mu\Delta t$:

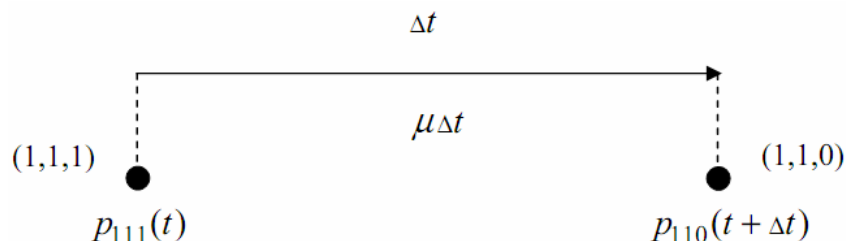


Рис. 5.38. Переход из состояния (1,1,1) в состояние (1,1,0) за время Δt

При переходе системы из состояния (1,1,1) в состояние (1,1,0) расчет вероятности $p_{110}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{110}(t + \Delta t) = p_{111}(t)\mu\Delta t. \quad (5.62)$$

Рассмотрим третий способ перехода в состояние (1,1,0). Вероятность того, что за время Δt не произойдет обслуживания требований ни в первой фазе, ни во второй фазе, определяется как $(1 - \mu\Delta t)(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.39) с учетом вероятности $p_{110}(t)$ и переходной вероятности $(1 - \mu\Delta t)(1 - \mu\Delta t)$:

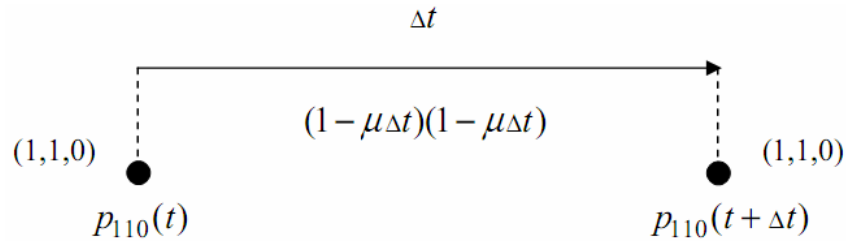


Рис. 5.39. Переход из состояния (1,1,0) в состояние (1,1,0) за время Δt

При переходе системы из состояния (1,1,0) в то же самое состояние (1,1,0) расчет вероятности $p_{110}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{110}(t + \Delta t) = p_{110}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.63)$$

Расчет вероятности $p_{110}(t + \Delta t)$ для всех трех способов перехода произведем по формуле полной вероятности с учетом (5.61), (5.62), (5.63):

$$p_{110}(t + \Delta t) = p_{010}(t)\lambda\Delta t(1 - \mu\Delta t) + p_{111}(t)\mu\Delta t + p_{110}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.64)$$

9. Рассмотрим состояние $(b,1,0)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{b10}(t + \Delta t)$. От

момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если произойдет обслуживание требования в первой фазе, а вторая фаза не завершит обслуживания: переход из состояния $(1,1,0)$ в состояние $(b,1,0)$;
- если в системе не произойдет изменение ее состояния: переход из состояния $(b,1,0)$ в то же самое состояние $(b,1,0)$;
- если произойдет обслуживание требования в третьей фазе и не произойдет обслуживания во второй фазе: переход из состояния $(b,1,1)$ в состояние $(b,1,0)$.

Рассмотрим первый способ перехода в состояние $(b,1,0)$. Вероятность того, что за время Δt в первой фазе произойдет обслуживание, а во второй фазе не произойдет, определяется как $\mu\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.40) с учетом вероятности $p_{110}(t)$ и переходной вероятности $\mu\Delta t(1 - \mu\Delta t)$:

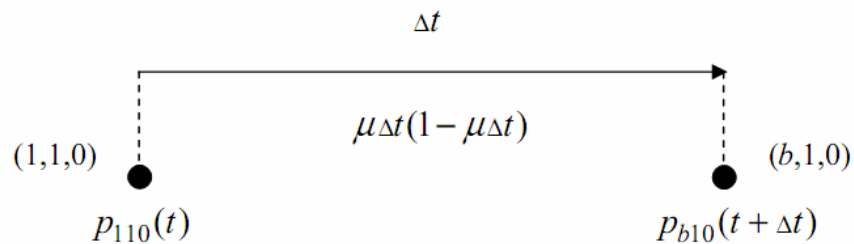


Рис. 5.40. Переход из состояния $(1,1,0)$ в состояние $(b,1,0)$ за время Δt

При переходе системы из состояния $(1,1,0)$ в состояние $(b,1,0)$ расчет вероятности $p_{b10}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{b10}(t + \Delta t) = p_{110}(t)\mu\Delta t(1 - \mu\Delta t). \quad (5.65)$$

Рассмотрим второй способ перехода в состояние $(b,1,0)$. Вероятность того, что за время Δt в системе не произойдет изменений, определяется

как $(1 - \mu\Delta t)$. Выполним графические пояснения второго способа перехода (рис. 5.41) с учетом вероятности $p_{b10}(t)$ и переходной вероятности $(1 - \mu\Delta t)$:

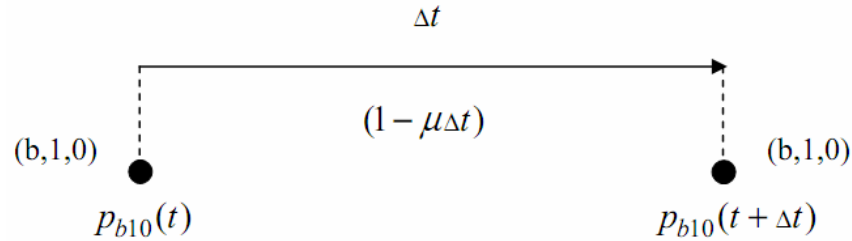


Рис. 5.41. Переход из состояния $(b,1,0)$ в состояние $(b,1,0)$ за время Δt

При переходе системы из состояния $(b,1,0)$ в то же самое состояние $(b,1,0)$ расчет $p_{b10}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{b10}(t + \Delta t) = p_{b10}(t)(1 - \mu\Delta t). \quad (5.66)$$

Рассмотрим третий способ перехода в состояние $(b,1,0)$. Вероятность того, что за время Δt в третьей фазе произойдет обслуживание требования, а во второй фазе не произойдет обслуживания, определяется как $(1 - \mu\Delta t)\mu\Delta t$. Выполним графические пояснения третьего способа перехода (рис. 5.42) с учетом вероятности $p_{b11}(t)$ и переходной вероятности $(1 - \mu\Delta t)\mu\Delta t$:

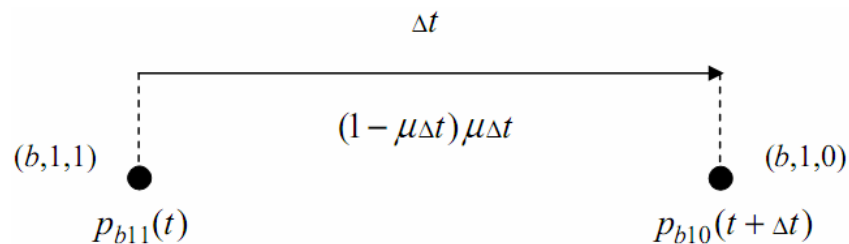


Рис. 5.42. Переход из состояния $(b,1,1)$ в состояние $(b,1,0)$ за время Δt

При переходе системы из состояния $(b,1,1)$ в состояние $(b,1,0)$ расчет вероятности $p_{b10}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{b10}(t + \Delta t) = p_{b11}(t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.67)$$

Расчет вероятности $p_{b10}(t + \Delta t)$ для всех трех способов перехода произведем по формуле полной вероятности с учетом (5.65), (5.66), (5.67):

$$p_{b10}(t + \Delta t) = p_{110}(t)\mu\Delta t(1 - \mu\Delta t) + p_{b10}(t)(1 - \mu\Delta t) + p_{b11}(t)(1 - \mu\Delta t)\mu\Delta t. \quad (5.68)$$

10. Рассмотрим состояние $(1, \mathbf{b}, 1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{1b1}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если не произойдет обслуживания требований в первой и третьей фазах и произойдет обслуживание во второй фазе: переход из состояния $(1, 1, 1)$ в состояние $(1, \mathbf{b}, 1)$;
- если не произойдет обслуживания требований в первой и третьей фазах: переход из состояния $(1, \mathbf{b}, 1)$ в то же самое состояние $(1, \mathbf{b}, 1)$;
- если в систему поступит требование, а в третьей фазе не произойдет обслуживания требования: переход из состояния $(0, \mathbf{b}, 1)$ в состояние $(1, \mathbf{b}, 1)$.

Рассмотрим первый способ перехода в состояние $(1, \mathbf{b}, 1)$. Вероятность того, что за время Δt в первой и третьей фазах не произойдет обслуживания требований, а во второй фазе обслужится одно требование, определяется как $(1 - \mu\Delta t)\mu\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.43) с учетом вероятности $p_{111}(t)$ и переходной вероятности $(1 - \mu\Delta t)\mu\Delta t(1 - \mu\Delta t)$:

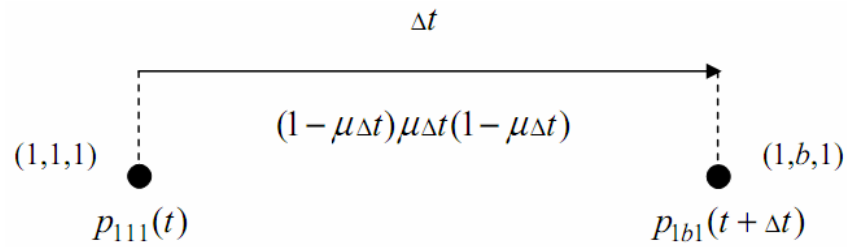


Рис. 5.43. Переход из состояния $(1,1,1)$ в состояние $(1,b,1)$ за время Δt

При переходе системы из состояния $(1,1,1)$ в состояние $(1,b,1)$ расчет вероятности $p_{1b1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{1b1}(t + \Delta t) = p_{111}(t)(1 - \mu\Delta t)\mu\Delta t(1 - \mu\Delta t). \quad (5.69)$$

Рассмотрим второй способ перехода в состояние $(1,b,1)$. Вероятность того, что за время Δt в первой и третьей фазах не произойдет обслуживание требований, определяется как $(1 - \mu\Delta t)(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.44) с учетом вероятности $p_{111}(t)$ и переходной вероятности $(1 - \mu\Delta t)(1 - \mu\Delta t)$:

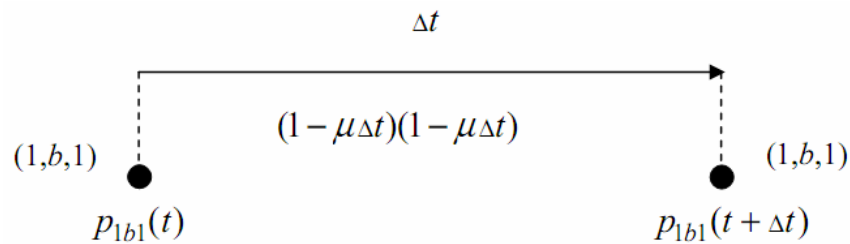


Рис. 5.44. Переход из состояния $(1,b,1)$ в состояние $(1,b,1)$ за время Δt

При переходе системы из состояния $(1,b,1)$ в то же самое состояние $(1,b,1)$ расчет вероятности $p_{1b1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{1b1}(t + \Delta t) = p_{1b1}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.70)$$

Рассмотрим третий способ перехода в состояние $(1,b,1)$. Вероятность того, что за время Δt в первую фазу поступит требование, а в третьей фазе не произойдет обслуживание требования, определяется как

$\lambda\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.45) с учетом вероятности $p_{0b1}(t)$ и переходной вероятности $\lambda\Delta t(1 - \mu\Delta t)$:

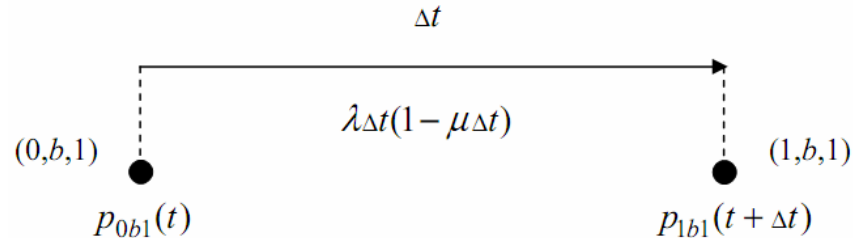


Рис. 5.45. Переход из состояния $(0, \mathbf{b}, 1)$ в состояние $(1, \mathbf{b}, 1)$ за время Δt
При переходе системы из состояния $(0, \mathbf{b}, 1)$ в состояние $(1, \mathbf{b}, 1)$ расчет вероятности $p_{1b1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{1b1}(t + \Delta t) = p_{0b1}(t)\lambda\Delta t(1 - \mu\Delta t). \quad (5.71)$$

Расчет вероятности $p_{1b1}(t + \Delta t)$ для всех трех способов перехода произведем по формуле полной вероятности с учетом (5.70), (5.72), (5.71):

$$p_{1b1}(t + \Delta t) = p_{111}(t)(1 - \mu\Delta t)^2\mu\Delta t + p_{1b1}(t)(1 - \mu\Delta t)^2 + p_{0b1}(t)\lambda\Delta t(1 - \mu\Delta t). \quad (5.72)$$

11. Рассмотрим состояние $(\mathbf{b}, \mathbf{b}, 1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{bb1}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти тремя способами:

- если в первой фазе произойдет обслуживание требования, а в третьей фазе обслуживания не произойдет: переход из состояния $(1, \mathbf{b}, 1)$ в состояние $(\mathbf{b}, \mathbf{b}, 1)$;
- если в системе не произойдет изменений: переход из состояния $(\mathbf{b}, \mathbf{b}, 1)$ в то же самое состояние $(\mathbf{b}, \mathbf{b}, 1)$ — когда не произойдет обслуживание требования в третьей фазе;

— если во второй фазе произойдет обслуживание требования, а в третьей фазе обслуживание требования не произойдет: переход системы из состояния $(b, 1, 1)$ в состояние $(b, b, 1)$.

Рассмотрим первый способ перехода в состояние $(b, b, 1)$. Вероятность того, что за время Δt в первой фазе произойдет обслуживание требования, а в третьей фазе не произойдет обслуживание требования, определяется как $\mu\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.46) с учетом вероятности $p_{1b1}(t)$ и переходной вероятности $\mu\Delta t(1 - \mu\Delta t)$:

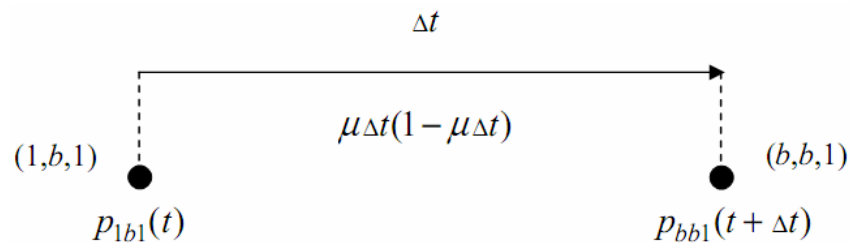


Рис. 5.46. Переход из состояния $(1, b, 1)$ в состояние $(b, b, 1)$ за время Δt

При переходе системы из состояния $(1, b, 1)$ в состояние $(b, b, 1)$ расчет вероятности $p_{bb1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{bb1}(t + \Delta t) = p_{1b1}(t)\mu\Delta t(1 - \mu\Delta t). \quad (5.73)$$

Рассмотрим второй способ перехода в состояние $(b, b, 1)$. Вероятность того, что за время Δt в системе не произойдет изменений — в третьей фазе не произойдет обслуживание требования, определяется как $(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.47) с учетом вероятности $p_{bb1}(t)$ и переходной вероятности $(1 - \mu\Delta t)$:

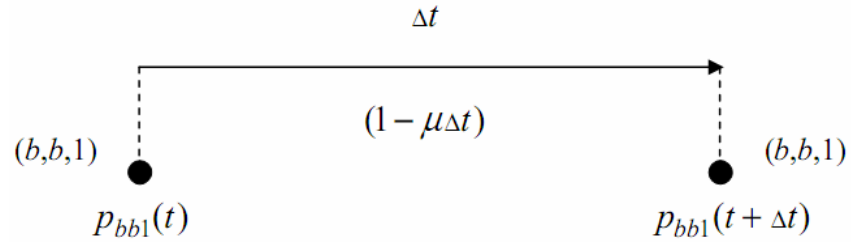


Рис. 5.47. переход из состояния $(b, b, 1)$ в состояние $(b, b, 1)$ за время Δt
При переходе системы из состояния $(b, b, 1)$ в то же самое состояние $(b, b, 1)$ расчет вероятности $p_{bb1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{bb1}(t + \Delta t) = p_{bb1}(t)(1 - \mu\Delta t). \quad (5.74)$$

Рассмотрим третий способ перехода в состояние $(b, b, 1)$. Вероятность того, что за время Δt во второй фазе произойдет обслуживание, а в третьей не произойдет, определяется как $\mu\Delta t(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения третьего способа перехода (рис. 5.48) с учетом вероятности $p_{b11}(t)$ и переходной вероятности $\mu\Delta t(1 - \mu\Delta t)$:

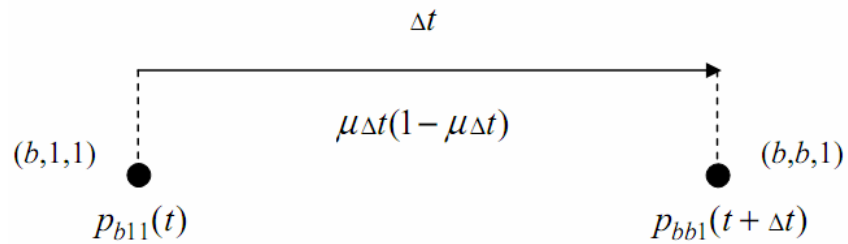


Рис. 5.48. Переход из состояния $(b, 1, 1)$ в состояние $(b, b, 1)$ за время Δt
При переходе системы из состояния $(b, 1, 1)$ в состояние $(b, b, 1)$ расчет вероятности $p_{bb1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{bb1}(t + \Delta t) = p_{b11}(t)\mu\Delta t(1 - \mu\Delta t). \quad (5.75)$$

Расчет вероятности $p_{bb1}(t + \Delta t)$ для всех трех способов перехода произведем по формуле полной вероятности с учетом (5.73), (5.74), (5.75):

$$p_{bb1}(t + \Delta t) = p_{1b1}(t)\mu\Delta t(1 - \mu\Delta t) + p_{bb1}(t)(1 - \mu\Delta t) + p_{b11}(t)\mu\Delta t(1 - \mu\Delta t). \quad (5.76)$$

12. Рассмотрим состояние $(b,1,1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{b11}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если первая фаза закончит обслуживание требования, а вторая и третья фазы не закончат: переход из состояния $(1,1,1)$ в состояние $(b,1,1)$;
- если в системе не произойдет изменений: переход из состояния $(b,1,1)$ в то же самое состояние $(b,1,1)$.

Рассмотрим первый способ перехода в состояние $(b,1,1)$. Вероятность того, что за время Δt в первой фазе произойдет обслуживание, а во второй и третьей фазах не произойдет, определяется как $\mu\Delta t(1 - \mu\Delta t)(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.49) с учетом вероятности $p_{111}(t)$ и переходной вероятности $\mu\Delta t(1 - \mu\Delta t)(1 - \mu\Delta t)$:

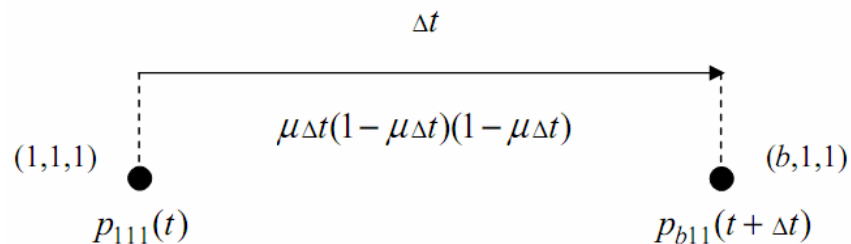


Рис. 5.49. Переход из состояния $(1,1,1)$ в состояние $(b,1,1)$ за время Δt

При переходе системы из состояния $(b,1,1)$ в состояние $(b,b,1)$ расчет вероятности $p_{b11}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{b11}(t + \Delta t) = p_{111}(t)\mu\Delta t(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.77)$$

Рассмотрим второй способ перехода в состояние $(b,1,1)$. Вероятность того, что за время Δt во второй и третьей фазах не произойдет

обслуживание требований, определяется как $(1 - \mu\Delta t)(1 - \mu\Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения второго способа перехода (рис. 5.50) с учетом вероятности $p_{b11}(t)$ и переходной вероятности $(1 - \mu\Delta t)(1 - \mu\Delta t)$:

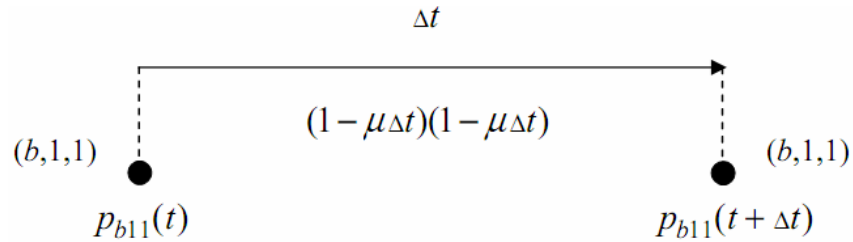


Рис. 5.50. Переход из состояния $(b,1,1)$ в состояние $(b,1,1)$ за время Δt

При переходе системы из состояния $(b,1,1)$ в состояние $(b,b,1)$ расчет вероятности $p_{b11}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{b11}(t + \Delta t) = p_{b11}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.78)$$

Расчет вероятности $p_{b11}(t + \Delta t)$ для обоих способов перехода произведем по формуле полной вероятности с учетом (5.77), (5.78):

$$p_{b11}(t + \Delta t) = p_{111}(t)\mu\Delta t(1 - \mu\Delta t)(1 - \mu\Delta t) + p_{b11}(t)(1 - \mu\Delta t)(1 - \mu\Delta t). \quad (5.79)$$

13. Рассмотрим состояние $(0,b,1)$ системы на момент времени $(t + \Delta t)$. Вероятность этого состояния обозначим, как $p_{0b1}(t + \Delta t)$. От момента времени t до момента времени $(t + \Delta t)$ в это состояние можно перейти двумя способами:

- если в систему не поступит требования, 2-я фаза завершит обслуживание требования, а 3-я фаза не закончит обслуживание требования: переход системы из состояния $(0,1,1)$ в состояние $(0,b,1)$;
- если система не изменит своего состояния: переход из состояния $(0,b,1)$ в то же самое состояние $(0,b,1)$.

Рассмотрим первый способ перехода в состояние $(0, \mathbf{b}, 1)$. Вероятность того, что за время Δt в первую фазу не поступит требование, а во второй фазе произойдет обслуживание требования и в третьей фазе не произойдет обслуживания требования, определяется как $(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t)$, которую назовем переходной вероятностью. Выполним графические пояснения первого способа перехода (рис. 5.51) с учетом вероятности $p_{011}(t)$ и переходной вероятности $(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t)$:

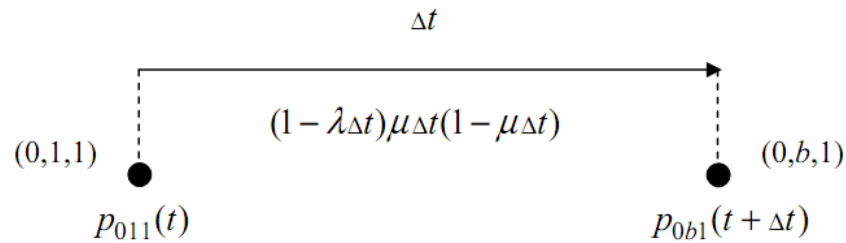


Рис. 5.51. Переход из состояния $(0, 1, 1)$ в состояние $(0, \mathbf{b}, 1)$ за время Δt

При переходе системы из состояния $(0, 1, 1)$ в состояние $(0, \mathbf{b}, 1)$ расчет вероятности $p_{0b1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{0b1}(t + \Delta t) = p_{011}(t)(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t). \quad (5.80)$$

Рассмотрим второй способ перехода в состояние $(0, \mathbf{b}, 1)$. Вероятность того, что за время Δt в первую фазу не поступит требование, а в третьей фазе не произойдет обслуживания требования, определяется как $(1 - \lambda \Delta t)(1 - \mu \Delta t)$. Выполним графические пояснения единственного способа перехода (рис. 5.52) с учетом вероятности $p_{0b1}(t)$ и переходной вероятности $(1 - \lambda \Delta t)(1 - \mu \Delta t)$:

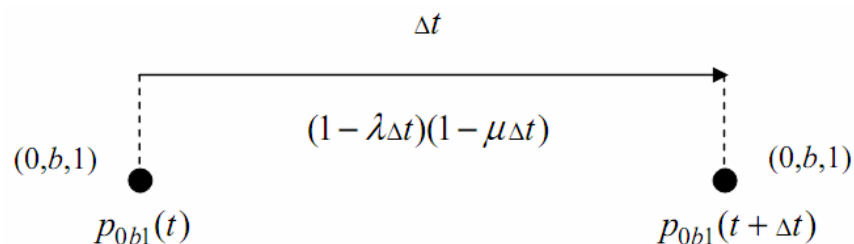


Рис. 5.39. Переход из состояния $(0, \mathbf{b}, 1)$ в состояние $(0, \mathbf{b}, 1)$ за время Δt
При переходе системы из состояния $(0, \mathbf{b}, 1)$ в то же самое состояние $(0, \mathbf{b}, 1)$ расчет вероятности $p_{0b1}(t + \Delta t)$ производится на основе теоремы о произведении вероятностей:

$$p_{0b1}(t + \Delta t) = p_{0b1}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t). \quad (5.81)$$

Расчет вероятности $p_{b11}(t + \Delta t)$ для обоих способов перехода произведем по формуле полной вероятности с учетом (5.80), (5.81):

$$p_{0b1}(t + \Delta t) = p_{011}(t)(1 - \lambda \Delta t)\mu \Delta t(1 - \mu \Delta t) + p_{0b1}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t). \quad (5.82)$$

Таким образом, получено 13 уравнений, связывающие вероятности состояний на момент времени $(t + \Delta t)$ с переходными вероятностями за время Δt и с вероятностями состояний на момент времени t . Выпишем эти уравнения:

$$\begin{aligned} p_{000}(t + \Delta t) &= p_{000}(t)(1 - \lambda \Delta t) + p_{001}(t)(1 - \lambda \Delta t)\mu \Delta t, \\ p_{100}(t + \Delta t) &= p_{000}(t)\lambda \Delta t + p_{100}(t)(1 - \mu \Delta t) + p_{101}(t)(1 - \mu \Delta t)\mu \Delta t, \\ p_{010}(t + \Delta t) &= p_{100}(t)(1 - \lambda \Delta t)\mu \Delta t + p_{010}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t) + p_{011}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t)\mu \Delta t, \\ p_{001}(t + \Delta t) &= p_{010}(t)(1 - \lambda \Delta t)\mu \Delta t + p_{001}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t) + p_{0b1}(t)(1 - \lambda \Delta t)\mu \Delta t, \\ p_{101}(t + \Delta t) &= p_{001}(t)\lambda \Delta t(1 - \mu \Delta t) + p_{101}(t)(1 - \mu \Delta t)(1 - \mu \Delta t) + p_{110}(t)\mu \Delta t + p_{1b1}(t)\mu \Delta t, \\ p_{011}(t + \Delta t) &= p_{101}(t)(1 - \lambda \Delta t)\mu \Delta t(1 - \mu \Delta t) + p_{011}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t)(1 - \mu \Delta t) + \\ &+ p_{b10}(1 - \lambda \Delta t)\mu \Delta t + p_{bb1}(t)(1 - \lambda \Delta t)\mu \Delta t, \\ p_{111}(t + \Delta t) &= p_{011}(t)\lambda \Delta t(1 - \mu \Delta t)(1 - \mu \Delta t) + p_{111}(t)(1 - \mu \Delta t)^3, \\ p_{110}(t + \Delta t) &= p_{010}(t)\lambda \Delta t(1 - \mu \Delta t) + p_{111}(t)\mu \Delta t + p_{110}(t)(1 - \mu \Delta t)(1 - \mu \Delta t), \\ p_{b10}(t + \Delta t) &= p_{110}(t)\mu \Delta t(1 - \mu \Delta t) + p_{b10}(t)(1 - \mu \Delta t) + p_{b11}(t)(1 - \mu \Delta t)\mu \Delta t, \\ p_{1b1}(t + \Delta t) &= p_{111}(t)(1 - \mu \Delta t)^2 \mu \Delta t + p_{1b1}(t)(1 - \mu \Delta t)^2 + p_{0b1}(t)\lambda \Delta t(1 - \mu \Delta t), \\ p_{bb1}(t + \Delta t) &= p_{1b1}(t)\mu \Delta t(1 - \mu \Delta t) + p_{bb1}(t)(1 - \mu \Delta t) + p_{b11}(t)\mu \Delta t(1 - \mu \Delta t), \\ p_{b11}(t + \Delta t) &= p_{111}(t)\mu \Delta t(1 - \mu \Delta t)(1 - \mu \Delta t) + p_{b11}(t)(1 - \mu \Delta t)(1 - \mu \Delta t), \\ p_{0b1}(t + \Delta t) &= p_{011}(t)(1 - \lambda \Delta t)\mu \Delta t(1 - \mu \Delta t) + p_{0b1}(t)(1 - \lambda \Delta t)(1 - \mu \Delta t). \end{aligned} \quad (5.83)$$

Преобразуем систему (5.83): раскроем скобки в правых частях, перенесем вероятности с индексами, соответствующими индексам вероятностям в левых частях, в левые части, обозначим произведение малых величин Δt как малую более высокого порядка, чем Δt , т. е. $o(\Delta t)$. Учтем также, что произведение малой величины $o(\Delta t)$ на постоянный множитель даст также малую величину, сумма двух малых величин также будет малая величина, произведение малых величин также малая величина. В итоге получим

$$\begin{aligned}
 p_{000}(t + \Delta t) - p_{000}(t) &= -p_{000}(t)\lambda\Delta t + p_{001}(t)\mu\Delta t + o(\Delta t), \\
 p_{100}(t + \Delta t) - p_{100}(t) &= p_{000}(t)\lambda\Delta t - p_{100}(t)\mu\Delta t + p_{101}(t)\mu\Delta t + o(\Delta t), \\
 p_{010}(t + \Delta t) - p_{010}(t) &= p_{100}(t)\mu\Delta t - p_{010}(t)(\lambda + \mu)\Delta t + p_{011}(t)\mu\Delta t + o(\Delta t), \\
 p_{001}(t + \Delta t) - p_{001}(t) &= p_{010}(t)\mu\Delta t - p_{001}(t)(\lambda + \mu)\Delta t + p_{0b1}(t)\mu\Delta t + o(\Delta t), \\
 p_{101}(t + \Delta t) - p_{101}(t) &= p_{001}(t)\lambda\Delta t - 2p_{101}(t)\mu\Delta t + p_{110}(t)\mu\Delta t + p_{1b1}(t)\mu\Delta t + o(\Delta t), \\
 p_{011}(t + \Delta t) - p_{011}(t) &= p_{101}(t)\mu\Delta t - p_{011}(t)(\lambda + 2\mu)\Delta t + p_{b10}\mu\Delta t + p_{bb1}(t)\mu\Delta t + o(\Delta t), \\
 p_{111}(t + \Delta t) - p_{111}(t) &= p_{011}(t)\lambda\Delta t - 3p_{111}(t)\mu\Delta t + o(\Delta t), \\
 p_{110}(t + \Delta t) - p_{110}(t) &= p_{010}(t)\lambda\Delta t + p_{111}(t)\mu\Delta t - 2p_{110}(t)\mu\Delta t + o(\Delta t), \\
 p_{b10}(t + \Delta t) - p_{b10}(t) &= p_{110}(t)\mu\Delta t - p_{b10}(t)\mu\Delta t + p_{b11}(t)\mu\Delta t + o(\Delta t), \\
 p_{1b1}(t + \Delta t) - p_{1b1}(t) &= p_{111}(t)\mu\Delta t - 2p_{1b1}(t)\mu\Delta t + p_{0b1}(t)\lambda\Delta t + o(\Delta t), \\
 p_{bb1}(t + \Delta t) - p_{bb1}(t) &= p_{1b1}(t)\mu\Delta t - p_{bb1}(t)\mu\Delta t + p_{b11}(t)\mu\Delta t + o(\Delta t), \\
 p_{b11}(t + \Delta t) - p_{b11}(t) &= p_{111}(t)\mu\Delta t - 2p_{b11}(t)\mu\Delta t + o(\Delta t), \\
 p_{0b1}(t + \Delta t) - p_{0b1}(t) &= p_{011}(t)\mu\Delta t - p_{0b1}(t)(\lambda + \mu)\Delta t + o(\Delta t).
 \end{aligned} \tag{5.84}$$

Обе части системы уравнений (5.84) разделим на Δt и перейдем к пределу от обеих частей уравнений при $\Delta t \rightarrow 0$. Учитывая, что

$$\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0, \text{ и, например, } \lim_{\Delta t \rightarrow 0} \frac{p_{000}(t + \Delta t) - p_{000}(t)}{\Delta t} = \frac{dp_{000}(t)}{dt},$$

получим следующую систему дифференциальных уравнений:

$$\begin{aligned}
 \frac{dp_{000}(t)}{dt} &= -\lambda p_{000}(t) + \mu p_{001}(t), \\
 \frac{dp_{100}(t)}{dt} &= \lambda p_{000}(t) - \mu p_{100}(t) + \mu p_{101}(t), \\
 \frac{dp_{010}(t)}{dt} &= \mu p_{100}(t) - (\lambda + \mu) p_{010}(t) + \mu p_{011}(t), \\
 \frac{dp_{001}(t)}{dt} &= \mu p_{010}(t) - (\lambda + \mu) p_{001}(t) + \mu p_{0b1}(t), \\
 \frac{dp_{101}(t)}{dt} &= \lambda p_{001}(t) - 2\mu p_{101}(t) + \mu p_{110}(t) + \mu p_{1b1}(t), \\
 \frac{dp_{011}(t)}{dt} &= \mu p_{101}(t) - (\lambda + 2\mu) p_{011}(t) + \mu p_{b10} + \mu p_{bb1}(t), \\
 \frac{dp_{111}(t)}{dt} &= \lambda p_{011}(t) - 3\mu p_{111}(t), \\
 \frac{dp_{110}(t)}{dt} &= \lambda p_{010}(t) + \mu p_{111}(t) - 2\mu p_{110}(t), \\
 \frac{dp_{b10}(t)}{dt} &= \mu p_{110}(t) - \mu p_{b10}(t) + \mu p_{b11}(t), \\
 \frac{dp_{1b1}(t)}{dt} &= \mu p_{111}(t) - 2\mu p_{1b1}(t) + \lambda p_{0b1}(t), \\
 \frac{dp_{bb1}(t)}{dt} &= \mu p_{1b1}(t) - \mu p_{bb1}(t) + \mu p_{b11}(t), \\
 \frac{dp_{b11}(t)}{dt} &= \mu p_{111}(t) - 2\mu p_{b11}(t), \\
 \frac{dp_{0b1}(t)}{dt} &= \mu p_{011}(t) - (\lambda + \mu) p_{0b1}(t).
 \end{aligned} \tag{5.85}$$

Полученная система (5.85) — это система однородных обыкновенных дифференциальных уравнений с постоянными коэффициентами (λ и μ). Для решения системы (5.85) следует задать начальные условия, которые могут быть естественными, т. е.

$$p_{000}(0)=1, \quad \forall p_{ijk}(0)=0, \quad i=0,1,b, \quad j=0,1,b, \quad k=0,1,b. \quad (5.86)$$

Таким образом, система дифференциальных уравнений (5.85) с начальными условиями (5.86) представляет собой математическую модель трехфазной системы массового обслуживания с нулевой вместимостью блоков ожидания.

5.3. Пример моделирования трехфазной системы массового обслуживания с нулевой вместимостью блоков ожидания

Представим систему (5.85) в матричном виде

$$\frac{dP}{dt} = AP, \quad (5.87)$$

где P — вектор вероятности состояний трехфазной системы размера 13×1 ,

A — матрица коэффициентов системы размера 13×13 .

Занумеруем вероятности состояний следующим образом:

$$\begin{aligned} p_{000}(t) &= p_1, & p_{100}(t) &= p_2, & p_{010}(t) &= p_3, & p_{001}(t) &= p_4, \\ p_{101}(t) &= p_5, & p_{011}(t) &= p_6, & p_{111}(t) &= p_7, & p_{110}(t) &= p_8, \\ p_{b10}(t) &= p_9, & p_{1b1}(t) &= p_{10}, & p_{bb1}(t) &= p_{11}, & p_{b11}(t) &= p_{12}, & p_{0b1}(t) &= p_{13}. \end{aligned} \quad (5.88)$$

Для введенной нумерации (5.88) перепишем систему дифференциальных уравнений (5.85) в виде:

$$\begin{aligned}\frac{dp_1(t)}{dt} &= -\lambda p_1(t) + \mu p_4(t), \\ \frac{dp_2(t)}{dt} &= \lambda p_1(t) - \mu p_2(t) + \mu p_5(t), \\ \frac{dp_3(t)}{dt} &= \mu p_2(t) - (\lambda + \mu) p_3(t) + \mu p_6(t), \\ \frac{dp_4(t)}{dt} &= \mu p_3(t) - (\lambda + \mu) p_4(t) + \mu p_{13}(t), \\ \frac{dp_5(t)}{dt} &= \lambda p_4(t) - 2\mu p_5(t) + \mu p_8(t) + \mu p_{10}(t), \\ \frac{dp_6(t)}{dt} &= \mu p_5(t) - (\lambda + 2\mu) p_6(t) + \mu p_9 + \mu p_{11}(t), \\ \frac{dp_7(t)}{dt} &= \lambda p_6(t) - 3\mu p_7(t), \\ \frac{dp_8(t)}{dt} &= \lambda p_3(t) + \mu p_7(t) - 2\mu p_8(t), \\ \frac{dp_9(t)}{dt} &= \mu p_8(t) - \mu p_9(t) + \mu p_{12}(t), \\ \frac{dp_{10}(t)}{dt} &= \mu p_7(t) - 2\mu p_{10}(t) + \lambda p_{13}(t), \\ \frac{dp_{11}(t)}{dt} &= \mu p_{10}(t) - \mu p_{11}(t) + \mu p_{12}(t), \\ \frac{dp_{12}(t)}{dt} &= \mu p_7(t) - 2\mu p_{12}(t), \\ \frac{dp_{13}(t)}{dt} &= \mu p_6(t) - (\lambda + \mu) p_{13}(t).\end{aligned}\tag{5.89}$$

Матрицу коэффициентов системы (5.89) будем обозначать, как A , и она принимает вид (приводится без обозначения):

$$\begin{bmatrix} -\lambda & 0 & 0 & \mu & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda & -\mu & 0 & 0 & \mu & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu & -(\lambda+\mu) & 0 & 0 & \mu & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu & -(\lambda+\mu) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu \\ 0 & 0 & 0 & \lambda & -2\mu & 0 & 0 & \mu & 0 & \mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & -(\lambda+2\mu) & 0 & 0 & \mu & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda & -3\mu & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 & \mu & -2\mu & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -\mu & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu & 0 & 0 & -2\mu & 0 & 0 & \lambda \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -\mu & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mu & 0 & 0 & 0 & 0 & -2\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu & 0 & 0 & 0 & 0 & 0 & 0 & -(\lambda+\mu) \end{bmatrix}.$$

Для интегрирования системы дифференциальных уравнений (5.89) зададим естественные начальные условия:

$$P_1(0) = 1, \quad P_2(0) = P_3(0) = \dots = P_{13}(0) = 0. \quad (5.90)$$

Программный код решения примера:

```
function three;  
clc,close  
format compact  
format long  
  
L = 3.3;  
M = 6.78;  
A = [-L, 0, 0, M, zeros(1,9);  
      L, -M, 0, 0, M, zeros(1,8);  
      0, M, -(L+M), 0, 0, M zeros(1,7);  
      0, 0, M, -(L+M), zeros(1,8), M;  
      0, 0, 0, L, -2*M, 0, 0, M, 0, M, 0, 0, 0;  
      zeros(1,4), M, -(L+2*M), 0, 0, M, 0, M, 0, 0;
```

```
zeros(1,5), L, -3*M, zeros(1,6);
0, 0, L, 0, 0, 0, M, -2*M zeros(1,5);
zeros(1,7), M, -M, 0, 0, M, 0;
zeros(1,6), M, 0, 0, -2*M, 0, 0, L;
zeros(1,9), M, -M, M, 0;
zeros(1,6), M, zeros(1,4), -2*M, 0;
zeros(1,5), M, zeros(1,6) -(L+M)];
P0 = [1;zeros(length(A)-1,1)];
T = [0, 4];
[t,P] = ode45(@fun3,T,P0, [], A);
Pycm = P(end,:)',
Control = sum(P(end,:))
plot(t,P,'linew',2.4),
grid on,

text(0.12,0.73,'\bf\fontsize{12}\leftarrow\it p\rm\bf_0(\itt\rm\bf)'),
ch = sprintf('%s%g%s%g', '\bf\it\fontsize{12}\lambda\rm\bf\fontsize{10}
= ', L,...
'\it\fontsize{12} \mu\rm\bf\fontsize{10} = ', M);
text(2.15,0.73, ch)
text(4.05,0,'\bf\it\fontsize{12} t'),
text(-0.2,1.1,'\bf\it\fontsize{12} P(t)'),

ylabel('\bf\fontsize{11} Вероятность '),
ylim([0, 1.05])
set(gcf,'color','w'),hold off

%% М-функция описания правой части дифференциальных
```

уравнений

function f = fun3(t, P, A)

f = A*P;

Диаграмма вероятностей, полученная в результате моделирования трехфазной системы, показана на рис. 5.40.

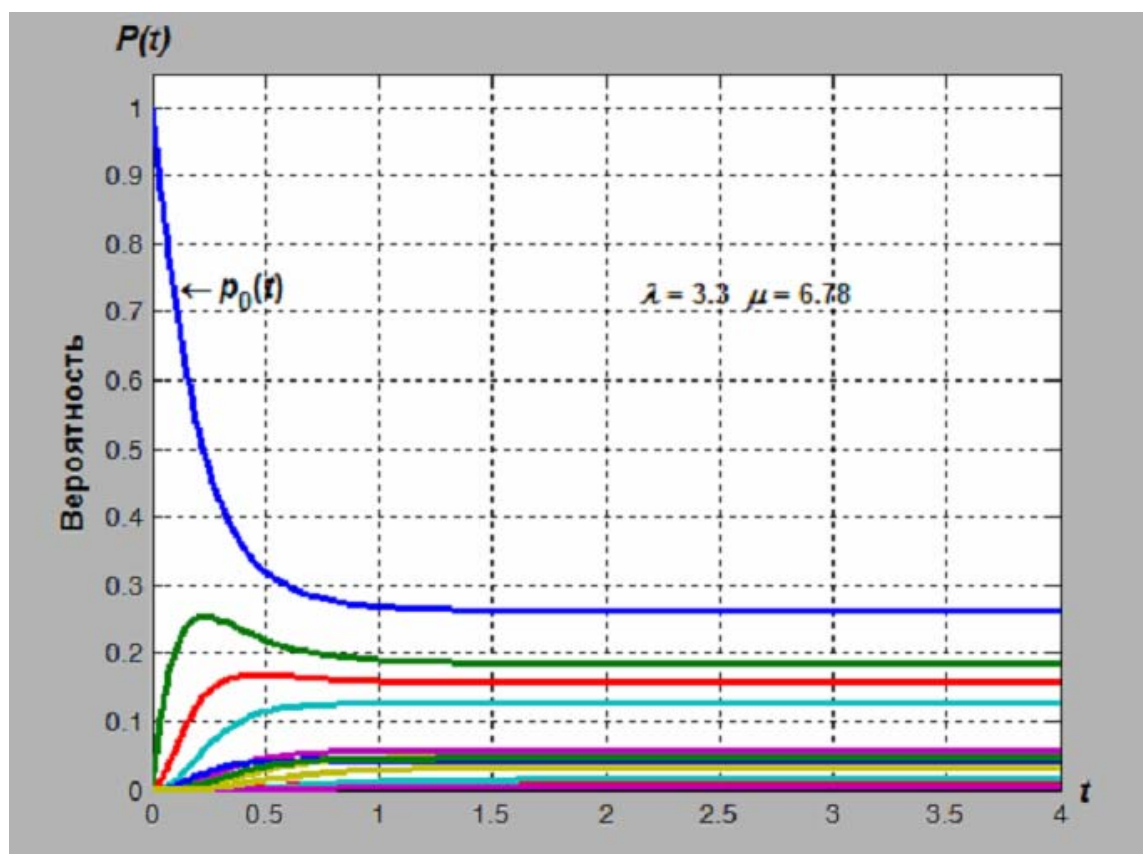


Рис. 5.40. Временные зависимости вероятностей трехфазной системы

По результату выполнения программы получены следующие установившиеся вероятности Руст системы и проверочное условие нормировки Control:

Руст =

0.262454243648217

0.185854450835115

0.157465066000584

0.127742446923626

0.058112724106246

0.048249759889706

0.007829320185467

0.042234368709798

0.046149784919119

0.011811741045176

0.015727075145721

0.003913653281267

0.032455365309959

Control =

1.0000000000000000

Произведем расчет основных операционных характеристик трехфазной системы.

1. Среднее число требований в системе — N_{cp} :

$$N_{cp} = \sum_{k=0}^3 k p_k,$$

где $p_k, k = \overline{0, 3}$ — финитные (стационарные вероятности), соответствующие тому, что в системе может находиться 0 требований, 1 требование, 2 требования, 3 требования. В соответствии с возможными состояниями системы, укажем вероятности, которые определяют состояния нахождения в системе 0, 1, 2, 3 требований.

В случае отсутствия требований вероятность обозначена как p_{000} , а с цифровым индексом по выражению (5.88) это будет p_1 .

Если в системе находится одно требование, то вероятности такого состояния обозначены как $p_{100}, p_{010}, p_{001}$. С цифровыми индексами по выражению (5.88) это будут p_2, p_3, p_4 .

Если в системе находятся два требования, то вероятности такого состояния обозначены как $p_{101}, p_{011}, p_{110}, p_{b10}, p_{0b1}$. С цифровыми индексами по выражению (5.88) это будут $p_5, p_6, p_8, p_9, p_{13}$.

Если в системе находятся три требования, то вероятности такого состояния обозначены как $p_{111}, p_{1b1}, p_{bb1}, p_{b11}$. С цифровыми индексами по выражению (5.88) это будут $p_7, p_{10}, p_{11}, p_{12}$.

Рассчитаем N_{cp} программным путем. Для этого вышеприведенную программу дополним следующими строками:

```
k = [0, 1, 2, 3];  
Pk = [P(end,1);sum(P(end,[2 3 4]));...  
sum(P(end,[5 6 8 9 13]));sum(P(end,[7 10:12]))];  
Ncp = k*Pk
```

Результат выполнения программы:

Ncp =
1.043311338601872

Таким образом, среднее число требований, находящихся в трехфазной системе массового обслуживания, примерно равно 1.0433. Очевидно, что среднее число требований в системе будет зависеть от соотношения интенсивности входного потока к интенсивности обслуживания.

2. Вероятность $P_{\phi 1}$ того, что вновь поступающее требование будет принято на обслуживание 1-й фазой:

$$P_{\phi 1} = p_{000} + p_{010} + p_{001} + p_{011} + p_{0b1}$$

или с цифровыми индексами по выражению (5.88):

$$P_{\phi 1} = p_1 + p_3 + p_4 + p_6 + p_{13}.$$

Рассчитаем $P_{\phi 1}$ по вышеприведенной основной программе с добавлением следующей строки:

$$\mathbf{Pf1 = sum(P(end,[1, 3, 4, 6, 13]))}$$

Результат вычисления:

Pf1 =

0.628366881772091

3. Вероятность отказа P_{omk} в обслуживании. Эта вероятность определяется как противоположная вероятность вероятности того, что поступающее требование будет принято на обслуживание 1-й фазой:

$$P_{omk} = 1 - P_{\phi 1} = 1 - 0.6284 = 0.3716.$$

4. Относительная пропускная способность Q рассматриваемой трехфазной системы по формальному определению:

$$Q = 1 - P_{omk} = 1 - 0.3716 = 0.6284.$$

Пропускная способность в данном случае равна вероятности $P_{\phi 1}$ того, что поступающее требование будет принято на обслуживание в 1-ю фазу.

5. Эффективная частота $\lambda_{эфф}$ поступления требований в трехфазную систему:

$$\lambda_{эфф} = \lambda P_{\phi 1}.$$

Для программного вычисления добавим следующую строку:

$$\mathbf{Lef = L*Pf1}$$

Результат вычисления:

Lef =

2.073610709847901

6. Абсолютная пропускная способность A данной трехфазной системы:

$$A = \lambda_{эфф} Q = 2.0736 \cdot 0.6284 = 1.30305024 \approx 1.3.$$

Абсолютная пропускная способность показывает среднее число требований, обслуженных в единицу времени данной трехфазной системой.

7. Полное время T_s пребывания требований (среднего числа) в трехфазной системе. Рассчитывается по формуле Литтла:

$$T_s = \frac{N_{cp}}{\lambda_{эфф}}.$$

Для программного вычисления добавим следующую строку:

$$Ts = Ncp/Lef$$

Результат вычисления:

$T_s =$

0.503137514504060

8. Среднее время T_{cp} обработки одного требования в трехфазной системе:

$$T_{cp} = 1/\mu + 1/\mu + 1/\mu = 3/\mu = 3/6.78 = 0.4425.$$

9. Средняя продолжительность задержки T_{bb} требования в 1-й и во 2-й фазах, когда они заблокированы (в условных единицах времени):

$$T_{bb} = T_s - T_{cp} = 0.5031 - 0.4425 = 0.0607.$$

10. Средняя продолжительность задержки T_b требования в одной из фаз (в 1-й или во 2-й), когда она заблокирована (в условных единицах времени):

$$T_b = \frac{T_{bb}}{2} = \frac{0.0607}{2} = 0.03035.$$

Рассмотренные операционные характеристики практически всесторонне позволяют оценить функционирование трехфазной системы массового обслуживания.

ГЛАВА 6. ПРИМЕНЕНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ В ЗАДАЧАХ АНАЛИЗА ПРОИЗВОДИТЕЛЬНОСТИ МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

6.1 Введение

В данной главе рассматриваются вопросы аналитического моделирования многопроцессорных SMP- и DSM- систем с разделяемой памятью и коммуникационной средой, выполненной на основе общей шины. Однако полученные результаты можно легко применить к структурам с другой коммуникационной средой (ПК, МШ), если внести соответствующие изменения в концептуальную модель коммутационной сети. Для MPP- систем адекватными являются модели систем с распределённой памятью, в которых должна быть изменена концептуальная модель передающей среды из коммутационной на связную.

В качестве инструмента моделирования используется теория стохастических сетей массового обслуживания, получившая в последнее время большое распространение и являющаяся частью общей теории массового обслуживания.

На системном этапе проектирования вычислительной системы проводят анализ различных вариантов структурной организации с целью получения их характеристик без построения реальной системы. Существуют два основных этапа анализа: макроанализ и микроанализ. Эти этапы отличаются не методами анализа (измерение, имитационное или аналитическое моделирование), а назначением.

Макроанализ производят на самом высоком уровне проектирования, когда определяется конфигурация системы, удовлетворяющая заданным требованиям. В качестве входных потоков здесь выступают потоки задач, поступающих в вычислительную систему. Результаты макроанализа используются для оценки общей производительности системы, времени ответа и получения других характеристик.

Микроанализ производят с целью получения характеристик вычислительной системы на более низких уровнях проектирования, т.е. на уровне отдельных устройств или частей системы. На этапе микроанализа более подробно описываются процессы, которые в макроанализе обычно игнорируются. В качестве входных потоков обычно выступают транзакции, формируемые исполняемыми командами. Результаты микроанализа чаще всего используют для оценки потерь производительности из-за аппаратных конфликтов в общих ресурсах (в блоках общей памяти, коммуникационной среде и др.), и связанную с ней оценку пропускной способности каналов обмена, блоков памяти и т.д.

Кроме количественных оценок макро- и микроанализ позволяет получать и качественные, поскольку в процессе аналитического моделирования вскрывается сущность процессов, происходящих в вычислительной системе. В этом заключается познавательная ценность метода аналитического моделирования.

Необходимость применения методов аналитического моделирования вычислительных систем объясняется существованием задержек вычислительного процесса из-за ограниченности ресурсов. Более всего этому явлению подвержены многопроцессорные системы, в которых ресурсы разделяются не только между устройствами, как в однопроцессорных ВС, но ещё и между несколькими процессорами (процессами), претендующими на доступ к ним.

Объяснить природу задержек или замедлений можно на следующем примере. Допустим, что на доступ к памяти претендуют процессор и канал прямого доступа в память. Каждый цикл, занятый каналом, оказывается циклом, который процессор занять уже не может. Если при отсутствии активности канала выполнение процесса требовалось N микросекунд, то каково будет полное время выполнения, если он занимает K циклов доступа к памяти? Очевидно, что число циклов памяти, требуемых для выполнения этого процесса, не зависит от активности канала, однако время, в течение которого осуществляются все эти обращения к памяти, возрастет в результате обслуживания канала. Если канал пассивен, то затраченное время и время выполнения будут одинаковы. Если же канал будет использовать память с максимально возможной для нее интенсивностью, то очевидно, что затраченное время возрастет до бесконечности. Таким образом, активность канала увеличивает время, затраченное процессом, т. е. замедляет его выполнение.

Если более приоритетные функции требуют своего обслуживания с *интенсивностью* R , а интенсивность доступа к памяти равна M , то потребность в памяти со стороны этих функций определяется как $p=R/M$. Если время обработки программы в отсутствие более приоритетных функций обозначить через T , то затраченное время $E=T/(1-p)$. Множитель $1/(1-p)$ называется коэффициентом замедления.

Замедление произойдет при следующих условиях:

- ресурс характеризуется максимальной интенсивностью обращений (например, число заданий на выполнение программ в секунду), которая является жестко фиксированной характеристикой аппаратуры. Чем выше интенсивность обращений к устройству, тем

более возрастает очередь к нему и, следовательно, увеличивается время ожидания обслуживания;

- ресурс, может использоваться более чем одним процессом одновременно, и существует по крайней мере один процесс, который может завладеть ресурсом раньше всех остальных, также претендующих на этот ресурс. Другой процесс при этом должен ждать освобождения ресурса.

Рассмотрим наиболее характерные источники задержек, типичные для большинства конфигураций аппаратных средств.

Соперничество за память. Это, вероятно, наиболее распространенная причина замедления. Память разделяется несколькими элементами, включая ЦП, каналы ПДП, другие каналы, а возможно, и некоторые другие элементы.

Соперничество за доступ к памяти среди потребляющих ее устройств системы может вызвать замедление, которое окажет влияние на разделение других ресурсов. Например, в случае прерывания соперничество за память связано также с соперничеством за ЦП. Здесь важно понять разницу между этими двумя соперничествами. Она не всегда очевидна, потому что быстродействие ЦП может соответствовать быстродействию памяти. К примеру, ЦП может выполнять 1 млн. опер./с, что будет соответствовать возможности памяти обеспечивать 2 млн. обращений в секунду.

Различие между соперничеством за ЦП и за память вызывается различием характеристик аппаратных средств:

- ЦП имеет достаточное количество регистров общего назначения, что позволяет избежать частых обращений к памяти, т. е. при обработке

будут преобладать межрегистровые операции. Так как быстродействия ЦП и памяти могут быть согласованы для наиболее неудачных сочетаний команд, то в ходе обычной обработки ЦП занимает далеко не все циклы обращения к памяти даже при отсутствии соперничества;

- ЦП использует для некоторых команд разделенные обращения к памяти (т. е. операции прочитать — изменить — записать), во время выполнения которых другие обращения к памяти блокируются;
- память физически организуется так, что каждое хранимое ею слово имеет длину, например, в один символ. ЦП выбирает коды команд и работает с операндами, имеющими длину в несколько символов. Быстродействие ЦП может оказаться ограниченным памятью и другими разделяющими ее процессами;
- память организуется в виде нескольких независимых блоков, причем порядок адресации таков, что логически последовательные символы находятся в разных ее блоках (память с расслоением). Это можно сделать, например, выбрав младшие биты логического адреса в качестве указателя номера блока. Пропускная способность памяти в настоящее время может быть выше, чем ЦП, поэтому имеется возможность осуществлять работу всех независимых блоков памяти одновременно без конфликтов (между разделяющими эти памяти устройствами) и поэтому без замедления;
- несколько ЦП системы разделяют независимо работающие памяти. Организация доступа такова, что эти ЦП обычно не влияют друг на друга, однако иногда такое влияние наблюдается.

При наличии одной или двух описанных ситуаций целесообразно принять простую систему доступа к памяти. При этом имеется возможность оценить коэффициент замедления достаточно точно

математически. Если же ситуация оказывается очень сложной и запутанной, как в случае больших многопроцессорных комплексов, то простые математические процедуры оказываются только грубо приближенными и, если требуется точность, следует обратиться к методам аналитического или имитационного моделирования.

Соперничество за ЦП. Основной причиной соперничества за ЦП являются внутренние и внешние прерывания, а также наличие нескольких программных состояний или уровней обработки, характеризующих мультипрограммный режим работы. Как уже отмечалось, процесс соперничества за ЦП неизбежно связан с процессом соперничества за память. Эти три понятия: внешние и внутренние прерывания и программные состояния часто путают, так как они тесно связаны во многих аппаратных реализациях. Однако эти понятия независимы, и влияние каждого из них на замедление в каждом случае может оказаться различным.

Об ЭВМ говорят, что она имеет несколько программных состояний, если, как минимум, в ней имеются средства записи состояния и хранения содержимого счетчика команд для каждого состояния, в котором она может функционировать. Такие средства могут быть крайне примитивными и сводиться только к признаку, показывающему состояние программы в данный момент (т. е. основной режим или фоновый режим), и команде, позволяющей сохранить значение счетчика команд (скажем, в стеке). И наоборот, каждому программному состоянию может быть отведен свой набор регистров. Изменение состояния может произойти в результате прерывания, обнаружения какого-либо условия либо посредством команды и приводит к потреблению, ресурсов, необходимых для сохранения старого состояния и установки нового (имеется в виду содержимое всех регистров и признаки условий). Обычно

во время изменения состояния обработка блокируется. В элементарной микро-ЭВМ, обладающей минимальными средствами изменения состояния, это изменение может потребовать нескольких сотен циклов обращения памяти и выполнения нескольких сотен команд. Если время пребывания в состоянии высокого приоритета значительно меньше (составляет, скажем, не более 5%) времени пребывания в состоянии низкого приоритета, то простой анализ замедления вполне осуществим. Если же смена состояний происходит часто (например, несколько тысяч раз за секунду) и (или) время пребывания в таких состояниях достаточно продолжительно (например, десятки миллисекунд), то упрощенные математические подходы не работают. И это снова приводит к необходимости использовать средства аналитического или имитационного моделирования, а иногда и прямых измерений.

Прерывания, посредством которых ЦП реагирует на внешние события, обычно приводят к изменению программного состояния. Прямо или косвенно прерывание подразумевает передачу управления новому состоянию, запоминание старого и инициирование программы, адрес начала которой выбирается в зависимости от источника прерывания. Механизмы прерывания могут быть достаточно сложными — прерывание вызывает изменение состояния и прямой переход на заранее определенную программу, которая предназначена для обработки только этого прерывания. И наоборот, механизмы могут быть элементарными - после запоминания старого значения счетчика команд осуществляется переход на общую для всех прерываний программу, которая программным путем устанавливает источник и причину прерывания. Так как изменение состояния является частью всего процесса обработки прерывания, то все замечания, относящиеся к анализу замедления изменения программного состояния, также применимы и к прерыванию.

Но так как прерывания предполагают дополнительную, обработку, то маловероятно, что прерывания могут допускать простое аналитическое представление.

Часто между "понятиями «внутренние» и «внешние» прерывания различий не делается, хотя в некоторых ЭВМ могут использоваться как те, так и другие. В этих случаях под внутренними прерываниями понимаются прерывания, возникающие в результате внутреннего обнаружения таких факторов, как ошибки, возникающие в АЛУ, обращения к запрещенным областям памяти, сбои в аппаратуре и т. д. При наличии в ЭВМ прерываний обоих типов механизмы их обработки, как правило, различны. Аналитически проблема здесь состоит в том, что ЭВМ может иметь несколько различных способов обработки прерываний, причем каждый подразумевает потребление различных ресурсов. Предположение, что все внутренние и внешние прерывания ведут себя одинаково, в конечном счете приведет к ошибке, поэтому необходимо выявить и проанализировать каждый механизм внутренних и внешних прерываний.

Соперничество за магистраль. Для упрощения анализа проблем, связанных с соперничеством за магистраль, целесообразно рассматривать магистраль как некий коммутатор, обеспечивающий связь составляющих систему устройств. Таким образом, вычислительную систему можно представить как совокупность устройств памяти, процессоров, каналов, и т. п. устройств, соединенных друг с другом средствами связи. Магистраль является простым разделяемым во времени переключателем $I \times N$. Возможны и более сложные структуры связей, включающие много магистралей и матричные коммутаторы.

Для большинства систем, содержащих только одну или две

магистральной, метод аналитического моделирования оказывается вполне приемлемым.

Разделение магистральной ведет к замедлению, т.к. при большой загрузке магистральной, например, в случае попытки одновременно занять магистраль, скажем, несколькими процессорами после обнаружения аномалии продолжится работа только одного устройства, другие будут приостановлены. Поэтому соперничество за магистраль может потребовать более детального анализа. Например, пусть имеется два устройства, например, процессоры, каждый из которых потребляет 25% пропускной способности магистральной. Можно было бы считать, что коэффициент использования магистральной процессорами составляет 50%, а это должно привести к увеличению в два раза задержки для всех обращающихся к этой магистральной процессов. Тем не менее это может быть верно только в первом приближении, а часто и совсем неверно. Последнее можно проиллюстрировать на следующих примерах:

- пересылки данных по магистральной могут происходить только последовательно. Таким образом, коэффициент замедления равен не 2, а только 1,33;
- мгновенное потребление равно 25% на процессор, однако цикл нагрузки составляет только 5%. Когда работает один процессор, замедление равно 1,33, если оба одновременно — замедление равно 2. Среднее замедление составит, однако, только 1,03;
- пересылка использует всю пропускную способность магистральной. Все остальные операции, связанные с обращениями к магистральной, на время пересылки приостанавливаются.

Уравнение для коэффициента замедления, данное ранее в этом разделе, подразумевает произвольную связь между причиной замедления

и его следствием. Если же на самом деле это не так, то может потребоваться более осторожное и тщательное исследование реальной ситуации, возможно, придется прибегнуть к блок-схеме процесса. Изучение ситуаций «соперничества» разделяющих ресурс процессов не ограничено лишь анализом магистрали, так как такие же ситуации могут возникнуть для каждого ресурса. Однако наиболее часто это случается при рассмотрении магистрали и реже — при рассмотрении ЦП.

Соперничество за канал. Если канал обслуживает только одно устройство, то проблемы соперничества не существует. Проблема может возникнуть только в отношении тех каналов, которые обслуживают несколько медленных асинхронных устройств. Анализ замедления при соперничестве за каналы подобен анализу магистрали. Отличие между ними носит скорее количественный, а не качественный характер. Передача информации по магистрали обычно происходит посимвольно или пословно. Следовательно, вполне возможны квазиодновременно выполняющиеся действия, которые могут привести к замедлению. Канал же, в отличие от магистрали, постоянно прикрепляется к устройству на все время обмена данными с ним.

Сложность анализа канала заключается в том, что данные передаются по каналу блоками, длина которых зависит от специфики устройства. Так, для высокоскоростного принтера — это строка, для дисплея — экран, для накопителя — страница и т. д. Знание того, каким образом данные передаются по каналу, весьма существенно для его анализа.

6.2. Основы теории стохастических сетей массового обслуживания

В этом подразделе приведены основные выкладки из теории систем массового обслуживания, которые будут применены в предложенных в учебном пособии моделях. Рассмотрены три вида систем массового обслуживания (СМО). К первому относятся классические СМО с неограниченной очередью типа $M/M/1$ и $M/M/m$ [1], причём первая буква M означает, что поток заявок на входе СМО является простейшим. Вторая буква M означает, что времена обслуживания в приборах распределены по экспоненциальному закону, третья показывает количество обслуживающих приборов в СМО. Второй вид – СМО типа $M/M/1/K$, где K означает конечность длины очереди. К третьему виду относятся СМО с приоритетным обслуживанием заявок. Кроме того здесь приведены основные параметры и характеристики стохастических сетей массового обслуживания.

6.2.1. Параметры стохастических сетей

Рассмотрим произвольную разомкнутую сеть массового обслуживания, содержащую n систем обслуживания (СМО) S_1, S_2, \dots, S_n , и единственный источник заявок, из которого они с вероятностью P_{0i} могут поступать в любую из СМО (рис.6.1).

Внутри сети заявки перемещаются случайным образом и могут покинуть её из любой СМО с вероятностью P_{i0} . Такие сети называются стохастическими.

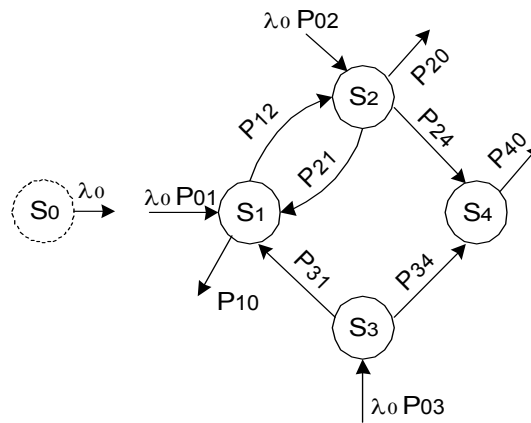


Рис. 6.1. Пример разомкнутой сети

Стохастическая сеть может быть задана следующими параметрами:

- числом n СМО в сети (S_0, S_1, \dots, S_n) , где S_0 – фиктивная СМО, моделирующая источник заявок, которая рассматривается как СМО с бесконечным числом заявок и с интенсивностью их обслуживания λ_0 ;
- числом каналов k_i в каждой СМО сети (k_1, k_2, \dots, k_n) ;
- интенсивностью потока заявок λ_0 источника S_0 ;
- интенсивностью потока на входе i -й СМО λ_i ;
- средним временем обслуживания заявок в каждой СМО сети (v_1, v_2, \dots, v_n) ;
- матрицей вероятности передач $P = [p_{ij}]$, где p_{ij} – вероятность того, что заявка покидающая СМО S_i , поступит на обслуживание в СМО S_j .

Для сети, состоящей из n СМО, матрица будет иметь следующий вид:

$$P = [p_{ij}] = \begin{matrix} & \begin{matrix} S_0 & S_1 & \dots & S_n \end{matrix} \\ \begin{matrix} S_0 \\ S_1 \\ \dots \\ S_n \end{matrix} & \begin{bmatrix} p_{00} & p_{01} & \dots & p_{0n} \\ p_{10} & p_{11} & \dots & p_{1n} \\ \dots & \dots & \dots & \dots \\ p_{n0} & p_{n1} & \dots & p_{nn} \end{bmatrix} \end{matrix}$$

Представленная матрица является стохастической. Для такой

матрицы должно выполняться условие $\sum_{j=0}^n p_{ij} = 1$, которое означает, что сумма элементов каждой строки матрицы равна единице.

Условие стохастической матрицы вытекает из того, что заявка, выходящая из S_i , обязательно попадает в некоторую СМО S_j .

Поскольку заявку из СМО S_i поступают в СМО S_j с вероятностью P_{ij} , то интенсивность потока, поступающего из S_i в S_j равна $\lambda_i p_{ij}$, где λ_i - интенсивность потока, выходящего из СМО S_i . Следовательно, с учётом всех источников на входе СМО S_j интенсивность будет определяться выражением

$$\lambda_j = \sum_{i=0}^n \lambda_i p_{ij}. \quad (6.1)$$

Коэффициент передачи в i -ю СМО $\alpha_i = \lambda_i / \lambda_0$ определяет число этапов обслуживания в i -й СМО, которое получит каждая заявка, поступающая на вход сети из источника моделируемого СМО.

6.2.2. Характеристики СМО с неограниченной длиной очереди.

Если заданы параметры сети, то можно определить следующие характеристики каждой СМО и сети в целом:

- среднюю длину очереди заявок в i -й СМО - l_i и в сети - L ;
- среднее число заявок, пребывающих в i -й СМО - m_i и в сети - M ;
- среднее время ожидания обслуживания заявки i -й СМО ω_i и в сети - W ;
- среднее время пребывания заявки в i -й СМО - u_i и в сети - U .

Состояние рассматриваемой сети описывается вектором

(a_1, a_2, \dots, a_n) , где a_i - число заявок в i -й СМО. Если через $p_i(a_i)$ обозначить распределение вероятностей того, что в установившемся

режиме в i -й СМО будет находится a_i заявок, то распределение числа заявок в сети определяется произведением распределений по всем СМО:

$$p(a_1, a_2, \dots, a_n) = p_1(a_1) p_2(a_2) \dots p_n(a_n), \quad (6.2)$$

где $p(a_i)$ - вероятность нахождения a_i заявок в i -й СМО.

Справедливость выражения (6.2) была доказана теоремой Джексона. Произведение распределений вероятностей в выражении (6.2) означает, что между отдельными СМО сети в значительной степени существует независимость. Это даёт основание декомпозировать сеть на n независимых СМО.

Если характеристики сети не меняются во времени (т.е. когда средние значения имеют постоянные величины), то в сети существует стационарный режим. Существование его связано с существованием стационарного режима в каждой СМО сети и определяется её загрузкой ρ_i . Для одноканальной СМО условием существования стационарного режима является:

$$\rho_i = \lambda_i v_i < 1 \quad (6.3)$$

для многоканальной -

$$\rho_i = \frac{\lambda_i v_i}{k_i} = \frac{\beta_i}{k_i} < 1, \quad (6.4)$$

где k_i - число каналов в i -й СМО; $\beta_i = \lambda_i v_i$ - среднее число занятых каналов. Стационарный режим означает, что в отдельных СМО и в сети в целом конечны средние задержки и длины очередей.

Рассмотрим характеристики отдельных СМО сети. Средняя длина очереди заявок для одноканальной СМО определяется выражением:

$$l_i = \frac{\rho_i^2}{1 - \rho_i}, \quad (6.5)$$

Для многоканальной СМО

$$l_i = \frac{\beta_i^{k_i+1}}{k_i! k_i \left(1 - \beta_i / k_i\right)^2} p_{0i}, \quad (6.6)$$

где p_{0i} - вероятность отсутствия заявок в i -й СМО,

$$p_{0i} = \left[\sum_{a_i=0}^{k_i-1} \frac{\beta_i^{a_i}}{a_i!} + \frac{\beta_i^{k_i}}{k_i! \left(1 - \beta_i / k_i\right)} \right]^{-1}$$

Среднее число заявок, пребывающих в системе, для одноканальной СМО определяется средней длиной очереди l_i и средним числом заявок, находящихся на обслуживании, ρ_i ,

$$m_i = l_i + \rho_i = \frac{\rho_i}{1 - \rho_i}. \quad (6.7)$$

Для многоканальной СМО среднее число заявок, находящихся на обслуживании, определяется числом занятых каналов, поэтому

$$m_i = l_i + \beta_i = \frac{\beta_i^{k_i+1}}{k_i! k_i \left(1 - \beta_i / k_i\right)^2} p_{0i} + \beta_i. \quad (6.8)$$

Среднее время ожидания заявки в очереди для одноканальной СМО равно частному от деления средней длины очереди l_i на интенсивность λ_i входящего в i -й СМО потока:

$$\omega_i = \frac{l_i}{\lambda_i} = \frac{v_i \cdot \rho_i}{1 - \rho_i}, \quad (6.9)$$

Для многоканальной СМО

$$\omega_i = \frac{l_i}{\lambda_i} = \frac{v_i \cdot \beta_i^{k_i}}{k_i! k_i \left(1 - \beta_i / k_i\right)^2} p_{0i}. \quad (6.10)$$

Среднее время пребывания заявки в системе определяется средней задержкой её в очереди и временем обслуживания в i -й СМО.

Для одноканальной СМО

$$u_i = \omega_i + v_i = \frac{v_i}{1 - \rho_1}; \quad (6.11)$$

для многоканальной

$$u_i = \omega_i + v_i = \frac{v_i \cdot \beta_i^{k_i}}{k_i! k_i \left(1 - \beta_i / k_i\right)^2} p_{0i} + v_i \quad (6.12)$$

На основании полученных характеристик отдельных СМО определяют характеристики сети в целом.

Среднее число заявок, ожидающих обслуживания в сети (т.е. среднее число заявок в очередях сети),

$$L = \sum_{i=1}^n l_i. \quad (6.13)$$

Среднее число заявок, пребывающих в сети,

$$M = \sum_{i=1}^n m_i \quad (6.14)$$

Поскольку каждая заявка может получить обслуживание в i -й СМО в среднем раз, то время ожидания обслуживания и время пребывания её в системе увеличится в раз. Среднее время ожидания заявки в очередях сети

$$W = \sum_{i=1}^n \alpha_i \omega_i, \quad (6.15)$$

а время пребывания

$$U = \sum_{i=1}^n \alpha_i u_i. \quad (6.16)$$

6.2.3 Характеристики СМО с ограниченной длиной очереди

Рассмотренные характеристики применимы для СМО с неограниченной длиной очереди. Однако в реальных системах число мест

в очереди может быть ограничено и заявка, застигнувшая очередь заполненной, получает отказ в обслуживании.

В случае ограниченной длины очереди перед одноканальной СМО с дисциплиной обслуживания FIFO, распределение вероятностей наличия в системе τ заявок определится выражением

$$P_{\tau i} = \frac{\rho_i^\tau (1 - \rho_i)}{1 - \rho_i^{R+1}} \quad \text{при } 0 \leq \tau \leq R, \quad (6.17)$$

где R – число мест в очереди обслуживания в i -й СМО.

Длина очереди к i -й СМО

$$l_i = \frac{1 - \rho_i}{1 - \rho_i^{R+1}} \left[\sum_{\tau=1}^R \tau \rho_i^\tau + 1 \right] - 1, \quad (6.18)$$

Число заявок, пребывающих в i -й СМО,

$$m_i = \frac{1 - \rho_i}{1 - \rho_i^{R+1}} \sum_{\tau=1}^R \tau \rho_i^\tau. \quad (6.19)$$

Время ожидания в очереди

$$\omega_i = \frac{1 - \rho_i}{\lambda_i (1 - \rho_i^{R+1})} \left[\sum_{\tau=1}^R \tau \rho_i^\tau + 1 \right] - \frac{1}{\lambda_i}. \quad (6.20)$$

Время пребывания в i -й СМО

$$u_i = \frac{1 - \rho_i}{\lambda_i (1 - \rho_i^{R+1})} \sum_{\tau=1}^R \tau \rho_i^\tau. \quad (6.21)$$

6.2.4. Характеристики СМО с приоритетами

Реальные системы не всегда используют дисциплину обслуживания *FIFO* (первая поступившая заявка обслуживается первой). Особенно часто встречающейся дисциплиной является приоритетная, в которой заявки, имеющие приоритет, обслуживаются вне очереди рис.(6.2).

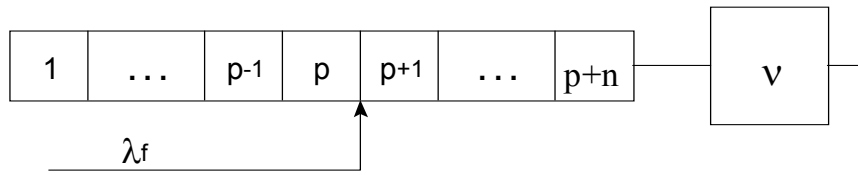


Рис. 6.2. СМО с приоритетным обслуживанием

Если заявка, имеющая приоритет, немедленно прерывает обслуживание другой заявки, то система работает с абсолютным приоритетом; если же заявка дожидается окончания обслуживания ранее поступившей, то система работает с относительным приоритетом.

Рассмотрим способ определения среднего времени ожидания для систем с относительным приоритетом.

Предположим, что поступающие заявки принадлежат одному из F различных приоритетных классов, обозначаемых через $f (f=1, 2, \dots, F)$. Будем считать, что чем больше значение индекса класса, тем выше приоритет этого класса. Таким образом, заявкам из приоритетного класса f представляется преимущественное обслуживание по сравнению с заявками из приоритетного класса $f-1$. Если в СМО находится несколько заявок с одинаковым приоритетом, то их обслуживание производится в соответствии с дисциплиной *FIFO*.

Рассмотрим модель, в которой заявки из приоритетного класса f образуют пуассоновский поток с интенсивностью λ_f , а время обслуживания ν_f распределено по экспоненциальному закону.

Определим параметры модели:

$$\lambda = \sum_{f=1}^F \lambda_f;$$

$$\nu = \sum_{f=1}^F \frac{\lambda_f}{\lambda} \nu_f;$$

$$p_f = \lambda_f \nu_f;$$

$$\rho = \lambda v = \sum_{f=1}^F \rho_f .$$

Здесь под ρ_f подразумевается загрузка обслуживающего прибора заявки из приоритетного класса f ; ρ - общая загрузка обслуживающего прибора, причём для обеспечения стационарного режима необходимо выполнение условия $\rho < 1$.

Введём следующие обозначения: ω_p - среднее время ожидания заявки из приоритетного класса f ; u_f - среднее время пребывания в СМО заявки, поступившей из приобретённого класса f .

Время ожидания разлагается на три составляющие:

- 1) время, связанное с тем, что в момент поступления данной заявки прибор был занят обслуживанием другой заявки;
- 2) время, связанное с тем, что в момент поступления данной заявки в очереди уже находятся заявки с приоритетом, равным или более высоким, чем у поступившей;
- 3) время, обусловленное тем, что заявки, поступившие позже, имеют более высокий приоритет.

Исследование приоритетных систем обычно начинают с поступившей вновь заявки из приоритетного класса f .

Будем называть эту заявку меченой. Обозначим через ω_0 среднюю задержку меченой заявки, связанную с наличием другой заявки на обслуживание; ω_0 будем определять через загрузку прибора ρ_i , которая представляет собой вероятность занятия обслуживающего прибора заявками из i -го класса. ρ_i также может быть интерпретирована как доля времени, в течение которого прибор занят заявками из i -го класса, тогда.

$$\omega_0 = \sum_{i=1}^F \rho_i . \quad (6.22)$$

Рассмотрим теперь вторую составляющую времени ожидания,

связанную с тем, что перед меченой заявкой обслуживаются заявки с равным или более высоким приоритетом, которые меченая заявка застала в очереди. Введём обозначение l_{if} - число заявок из класса i , которые застала в очереди меченая заявка из класса f . В соответствии с формулой Литтла $l_{if} = \lambda_i \omega_i$, где $i=f, f+1, \dots, F$. Отсюда средняя задержка меченой заявки составит

$$\sum_{i=f}^F v_i \lambda_i \omega_i .$$

Аналогично можно определить третью составляющую среднего времени ожидания (задержка меченой заявки за счёт того, что заявки, поступающие после неё, имеют более высокий приоритет). Пусть m_{if} - число заявок из класса i , поступающих в СМО, когда меченая заявка (из класса f) находится в очереди, и получающих обслуживание раньше меченой заявки.

Тогда $m_{if} = \lambda_i \omega_{if}$, где $i=f+1, f+2, \dots, F$, а средняя задержка составит

$$\sum_{i=f+1}^F \lambda_i \omega_f .$$

Таким образом, для СМО с относительным приоритетом

$$\omega_f = \omega_o + \sum_{i=f}^F v_i \lambda_i \omega_i + \sum_{i=f+1}^F v_i \lambda_i \omega_f . \quad (6.23)$$

Так как $v_i \lambda_i = \rho_i$, то, преобразовав выражение относительно ω_f , получим

$$\omega_f = \frac{\omega_o + \sum_{i=f+1}^F v_i \lambda_i \omega_f}{1 - \sum_{i=f}^F \rho_i} , \quad (6.24)$$

где $f=1, 2, \dots, F$

Система уравнений (6.24) может быть легко решена рекуррентно, т.е. сначала находится ω_F , затем ω_{F-1} и т.д.

Например, имеется три класса приоритетов, причём наибольший приоритет имеет класс 3. Тогда среднее время задержки каждого класса определится так:

$$\begin{aligned}\omega_3 &= \frac{\omega_o}{1 - \rho_3}; \\ \omega_2 &= \frac{\omega_o + \rho_3 \omega_3}{1 - \rho_2 - \rho_3}; \\ \omega_1 &= \frac{\omega_o + \rho_2 \omega_2 + \rho_3 \omega_3}{1 - \rho_1 - \rho_2 - \rho_3}.\end{aligned}$$

6.3. Макроанализ производительности многопроцессорных систем

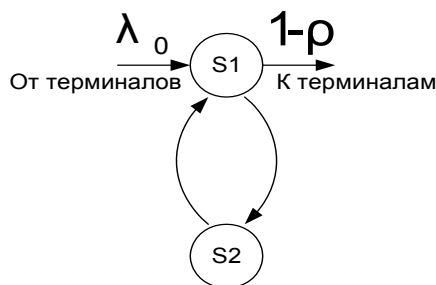
6.3.1. Модель с разделяемой памятью

Многопроцессорная система с разделяемой (общей) памятью приведена на рис. 6.2. Предварительно для упрощения будем считать, что в системе отсутствуют как внутрипроцессорные, так и внешние кэш. Виртуальная память системы содержит два уровня, причём на первом уровне находится основная память, являющаяся общей для всех процессорных модулей, а на втором уровне – ВЗУ. Обмен между уровнями производится посредством канала прямого доступа.

Предлагается, что в основной памяти системы размещается только часть программ и данных, используемых оперативно, другая часть программ и данных храниться в ВЗУ.

Будем считать, что система однородна, т.е. все процессоры одинаковы. Однородность предоставляет возможность системе исполнять любую программу на любом процессоре. Поскольку любой процессор может принять на обслуживание любую заявку, то система работает в режиме разделения нагрузки.

Будем считать, что задания на решение задач формируются с помощью терминалов (на рисунке ниже терминалы представлены как устройства ввода/вывода - УВВ), т.е. поступают в систему извне. В процессе решения задача с некоторой вероятностью p может обращаться в ВЗУ за недостающими программами или данными (см. рис.6.1). Обслуженные в ВЗУ задачи возвращаются на дообслуживание в процессор. Решенные задачи с вероятностью $(1-p)$ выходят из системы и возвращаются на терминалы пользователей. Таким образом, рассматриваемая многопроцессорная система функционирует в режиме оперативной обработки.



Функционирование многопроцессорной системы с двухуровневой памятью в режиме оперативной обработки можно представить в виде разомкнутой сетевой модели. Такая модель позволяет проанализировать задержки, связанные с обращением задач в ВЗУ и с ограничением числа процессоров в системе. Чтобы не усложнять модель детализацией, будем игнорировать потери, связанные с ограничением других ресурсов.

Будем считать, что на вход сети поступает пуассоновский поток заявок с интенсивностью λ_0 . Для реальной системы это соответствует потоку заданий на входе многопроцессорной системы. Кроме того принято, что времена обслуживания во всех устройствах распределены по экспоненциальному закону. Заявка обслуживается в начале в СМО S_1 , которая моделирует обработку, происходящую в системе процессор - память (ЦП/ОП). Получив обслуживание в S_1 , она с некоторой

вероятностью p обращается в СМО, которая моделирует ВЗУ (S_2). С вероятностью $(1-p)$ заявка покидает систему.

Модель системы с разделяемой (общей) памятью, одним каналом ПДП и k индивидуальными ВЗУ представлена на рис.6.2.

Поскольку каждый процессор может выполнять любую заявку, размещённую в общей памяти, то очередь для всех процессоров единая. Модель процесса выполнения заявок в процессорах может быть представлена многоканальной СМО S_1 с временем обслуживания ν_1 .

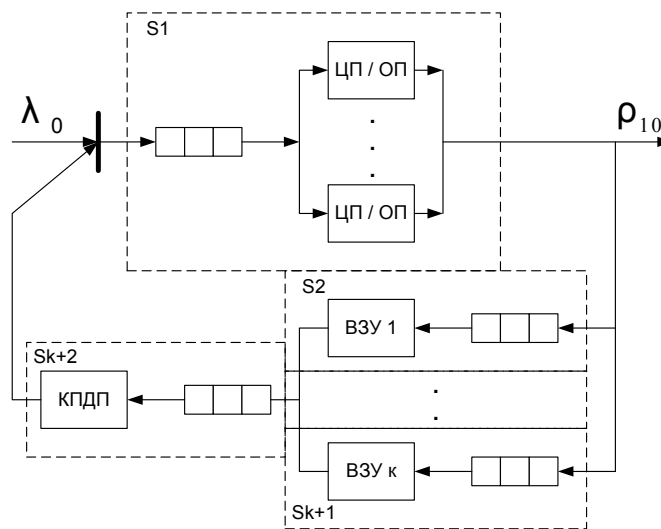


Рис. 6.2. Модель системы с разделяемой памятью и индивидуальными ВЗУ

Вначале заявка получает обслуживание в СМО S_1 . Если процесс её обработки закончился, то она с вероятностью p_{10} покидает систему. Если задаче в процессе её обработки потребуется обращение к ВЗУ, то она с вероятностью p_{1i} (где $i=2, \dots, k+1$) поступает на обслуживание в СМО S_i , имеющее среднее время обслуживания ν_i . Если заявка допускает обслуживание только в одном ВЗУ, то система внешней памяти представляется совокупностью одноканальных СМО

Поскольку весь процесс обмена с внешней памятью ведётся через канал прямого доступа, то модель КПДП может быть представлена как

одноканальная СМО с временем обслуживания ν_{k+2} .

Таким образом модель, изображенная на рис.6.2. представляется как совокупность одноканальных и многоканальных СМО, объединённых в стохастическую сеть.

Граф передач такой сети будет иметь вид, изображённый на рис.6.3. Над дугами указаны вероятности передач из одной СМО в другую. Над теми дугами, где вероятности не указаны, их значения равны единице.

Матрица вероятностей передач для такой сети будет иметь вид:

$$P = [p_{ij}] = \begin{matrix} & \begin{matrix} S_0 & S_1 & S_2 & \dots & S_{k+2} \end{matrix} \\ \begin{matrix} S_0 \\ S_1 \\ S_2 \\ \dots \\ S_{k+2} \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ p_{10} & 0 & p_{12} & \dots & p_{1,k+2} \\ 0 & 0 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \end{matrix}$$

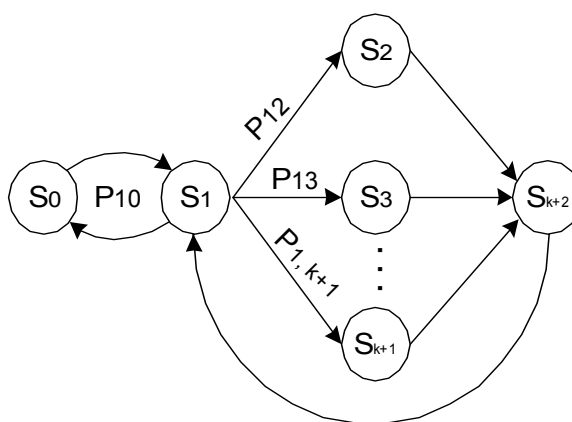


Рис. 6.3. Граф передач сети, изображенной на рис. 6.2

Для определения характеристик отдельных характеристик отдельных СМО и всей сети в целом необходимо знать интенсивности потоков заявок на входе каждой СМО. Для этого можно воспользоваться системой уравнений (2.1), которая будет состоять из $k+2$ уравнений:

$$\begin{aligned}\lambda_0 &= \lambda_1 \cdot P_{10}; \\ \lambda_1 &= \lambda_0 + \lambda_{k+2}; \\ \lambda_2 &= \lambda_1 \cdot P_{12}; \\ &\dots \\ \lambda_{k+1} &= \lambda_1 \cdot P_{1,k=1}; \\ \lambda_{k+2} &= \lambda_2 + \lambda_3 + \dots + \lambda_{k+1}\end{aligned}$$

Если в сети существует стационарный режим, то заявка, поступившая во внешнюю память, может обслуживаться любым ВЗУ и модель внешней памяти будет представлена в виде многоканальной СМО (рис.6.4) с временем обслуживания v_2 .

На рис.6.5 показан граф передач такой сети. Параметры и характеристики такой сети определяются аналогично предыдущей модели.

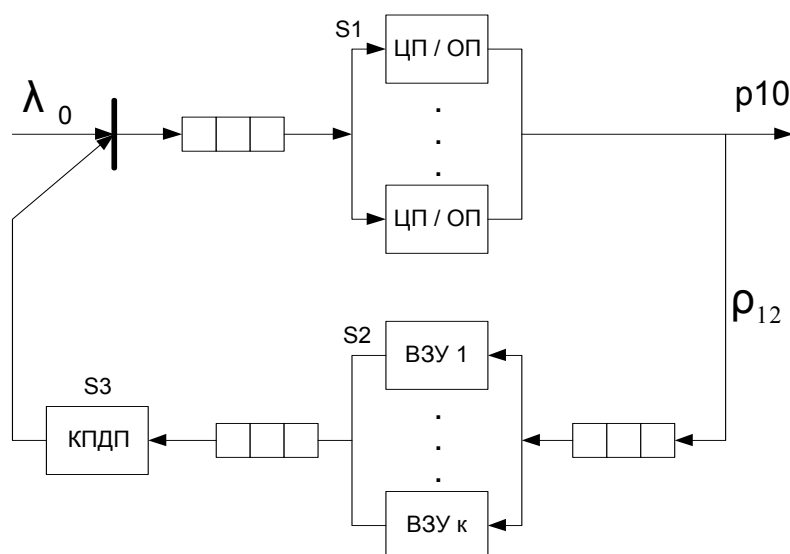


Рис. 6.4. Модель системы с разделяемой памятью и общими ВЗУ

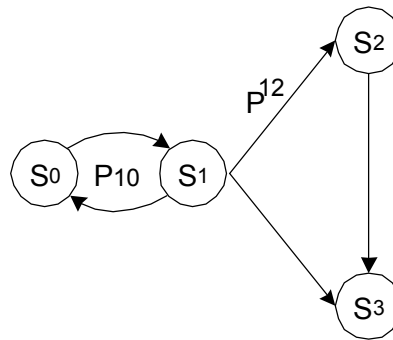


Рис. 6.5. Граф передач сети, изображенной на рис. 6.4

6.3.2. Модель системы с разделяемой распределённой памятью

Структура многопроцессорной системы с разделяемой распределённой памятью изображена на рис. 6.2. Виртуальная память системы, как и в предыдущем случае, является двухуровневой, состоящей из локальной памяти, размещенной в процессорных модулях, и ВЗУ. Кэш-память с целью упрощения модели не представлена ни на одном уровне.

В такой системе каждый процессор в основном работает с локальной памятью, в которой хранятся копии необходимых для функционирования программ и данных, загружаемых из внешней памяти. В процессе решения задач процессоры могут обращаться во внешнюю память за недостающими программами и данными.

Поскольку процессоры работают в основном самостоятельно (такие системы называют с разделением функций), то их можно представить в модели как n отдельных СМО с временем обслуживания $\nu_1, \nu_2, \dots, \nu_n$.

В зависимости от способа обслуживания внешней памятью возможны два варианта моделей. Если заявка обслуживается только в одном ВЗУ, то модель системы примет вид, изображённый на рис.6.6.

Она представляется как совокупность отдельных СМО ($S_1, S_2, \dots, S_{n+1}, \dots, S_{n+k}, S_{n+k+1}$), объединённых в стохастическую сеть. Граф передач

такой сети будет иметь вид, изображённый на рис.6.7.

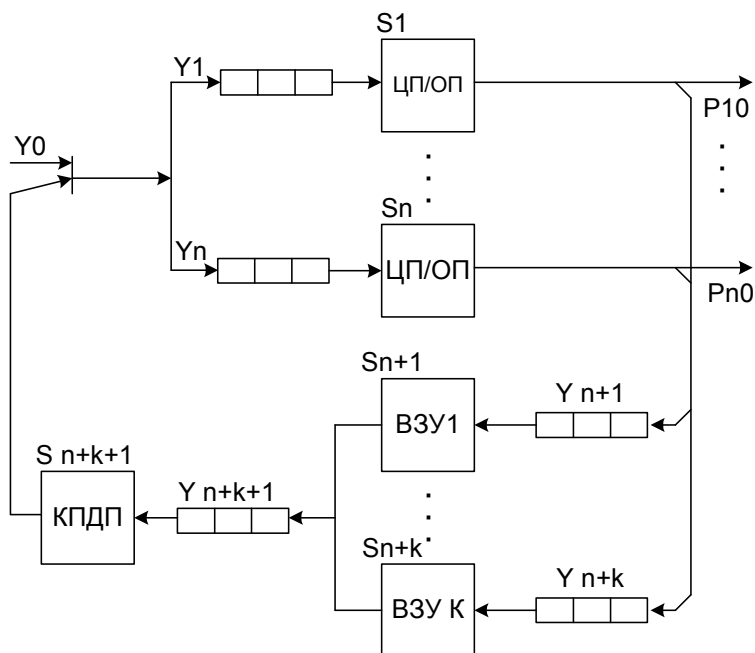


Рис.6.6. Модель системы с разделяемой распределённой памятью и индивидуальными ВЗУ.

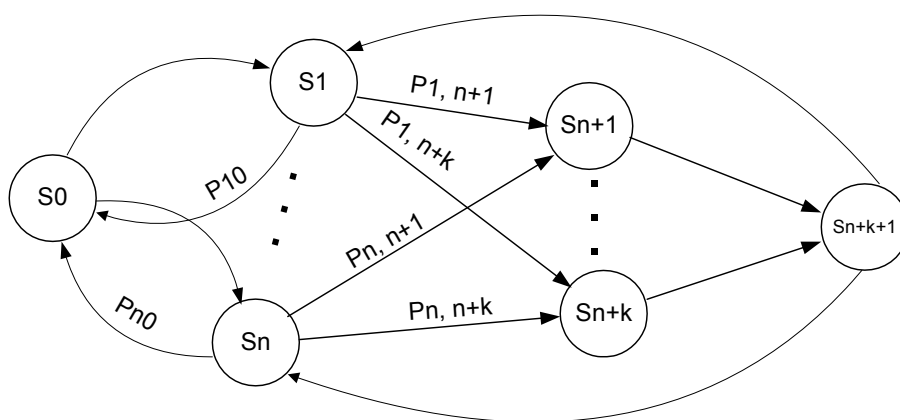


Рис.6.7. Граф передач сети, изображённой на рис.6.6.

Интенсивность потоков на входах СМО определяется из системы уравнений (6.1):

$$\lambda_o = \lambda_1 P_{10} + \lambda_2 P_{20} + \dots + \lambda_n P_{n0};$$

$$\lambda_1 = \lambda_o + \lambda_{n+k+1} P_{n+k+1,1};$$

$$\lambda_2 = \lambda_o + \lambda_{n+k+1} P_{n+k+1,2};$$

...

$$\lambda_n = \lambda_o + \lambda_{n+k+1} P_{n+k+1,n};$$

...

$$\lambda_{n+k} = \lambda_1 + P_{1,n+k} + \lambda_2 P_{2,n+k} + \dots + \lambda_n P_{n,n+k};$$

$$\lambda_{n+k+1} = \lambda_1 + \lambda_2 + \dots + \lambda_n;$$

$$\lambda_{n+1} = \lambda_1 + P_{1,n+1} + \lambda_2 P_{2,n+1} + \dots + \lambda_n P_{n,n+1};$$

Стационарный режим в сети будет существовать в том случае, когда выполняется условие (6.3). Характеристики отдельных СМО и сети в целом определяются выражениями (6.5), (6.7), (6.9), (6.11), (6.13)-(6.16).

Если заявка, поступившая во внешнюю память, может обслуживаться любым ВЗУ, то её модель будет многоканальной СМО. Тогда модель всей системы примет вид, изображённый на рис.6.7. Граф передач сети будет иметь вид, изображённый на рис.6.8.

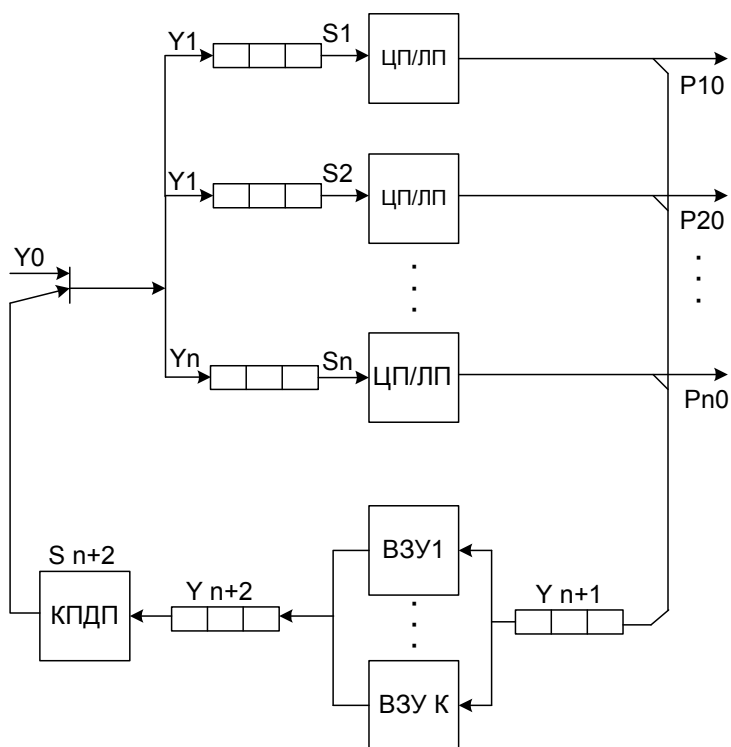


Рис.6.7. Модель системы с разделяемой распределённой памятью и общими ВЗУ.

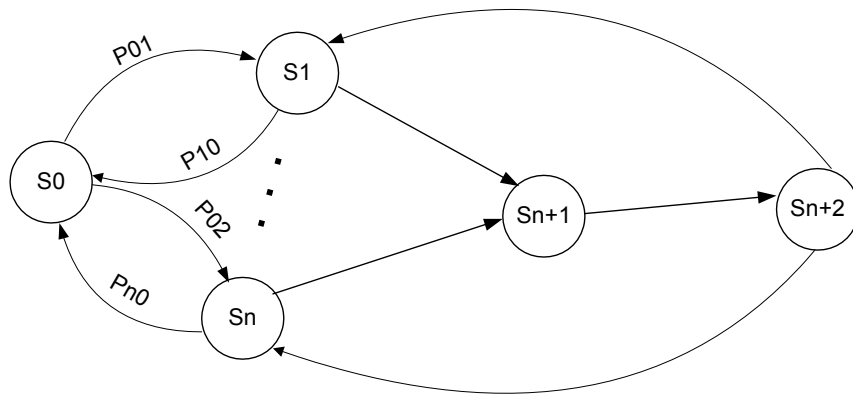


Рис.6.8. Граф передач сети, изображённой на рис.6.7.

Параметры и характеристики отдельных СМО и сети в целом определяются по выражениям (6.1), (6.3)-(6.16).

6.4. Микроанализ производительности многопроцессорных систем

Микроанализ используют в качестве инструмента исследования ВС на более низких уровнях проектирования по сравнению с макроанализом. В его задачу может входить оценка пропускной способности коммуникационной среды, оценка влияния различных методов доступа к общей памяти (например, по прерываниям или прямым доступом) на её пропускную способность, оценку времени доступа к файлам, размещаемым в ВЗУ, а также оценка задержек в обслуживании, связанных с аппаратными конфликтами, возникающими в общих ресурсах. Но чаще всего микроанализ производят с целью оценки потерь производительности. Почему необходима такая оценка? Дело в том, что производительность многопроцессорной системы всегда ниже суммарной производительности входящих в неё процессоров. Это объясняется в основном конечностью числа ресурсов системы, находящихся в общем пользовании. Конечность числа ресурсов приводит к очередям, возникающим при запросе этих ресурсов, а следовательно, к задержкам в обслуживании.

В системах с ОШ в качестве общих ресурсов могут выступать буферные ЗУ, используемые в каналах передачи данных, блоки разделяемой памяти, внешние запоминающие устройства, шина обмена.

В дальнейшем будут рассмотрены оценки влияния аппаратных конфликтов, возникающих в общей шине и общей памяти на производительность многопроцессорных систем с ОШ. Для этого необходимо более подробно рассмотреть структурную организацию центральной части таких систем.

6.4.1. Модель системы с разделяемой памятью для анализа конфликтов, возникающих в общих ресурсах

Модель представляется в виде разомкнутой стохастической сети, в которой источником заявок выступают процессоры, генерирующие поток команд (в отличие от моделей для макроанализа, где источником заявок являются задания, поступающие на обработку в ВС извне, например, с терминалов), а в качестве обслуживающих приборов – ОШ и модули общей памяти.

Модель источника заявок может быть представлена так, как изображено на рис.6.9. φ_i моделирует поток команд обращения i -го процессора к общей памяти, γ_i - к локальной памяти. В структурах систем с общей памятью поток γ_i отсутствует, т.к. отсутствует локальная память; в структурах с локальной памятью действуют оба потока.

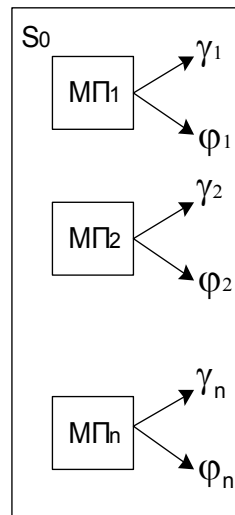


Рис.6.9. Модель источника заявок.

Процесс выполнения заявок из потока φ_i проходит две фазы обслуживания: вначале – общей шиной S_1 , затем одним из модулей памяти $S_j (j=2, \dots, m+1)$. Заявки из потока γ_i проходят одну фазу – только модуля памяти S_i . Будем считать все потоки простейшими, а времена обслуживания распределёнными по экспоненциальному закону.

Структурная схема модели системы с общей памятью показана на рис.6.17. Интенсивность потока заявок на входе сети определяется суммарным потоком заявок φ_i процессоров. Заявка процессора, поступившая в СМО S_1 и заставшая его занятым, становится в очередь, которую будем считать неограниченной. При обслуживании с приоритетами очередь перед СМО S_1 устанавливается в соответствии с классом приоритета заявки.

Заявка, обслуженная в СМО S_1 , с вероятностью p_{ij} поступает на обслуживание в одну из СМО $S_j (j=2, \dots, m+1)$. Процесс обслуживания в СМО S_j может быть либо в соответствии с дисциплиной *FIFO*, либо с приоритетной дисциплиной. Если очередь перед СМО S_2, \dots, S_{m+1} ограничена, то в случае её занятости заявка получит отказ в обслуживании и с вероятностью p_{11} возвратится в конец очереди к СМО S_1 (при беспriorитетной дисциплине обслуживания). В случае

приоритетного обслуживания СМО S_1 заявка, получившая отказ в обслуживании в одной из СМО S_j , приобретает тот же приоритет, который имела ранее. Получив полное обслуживание, заявка покидает сеть.

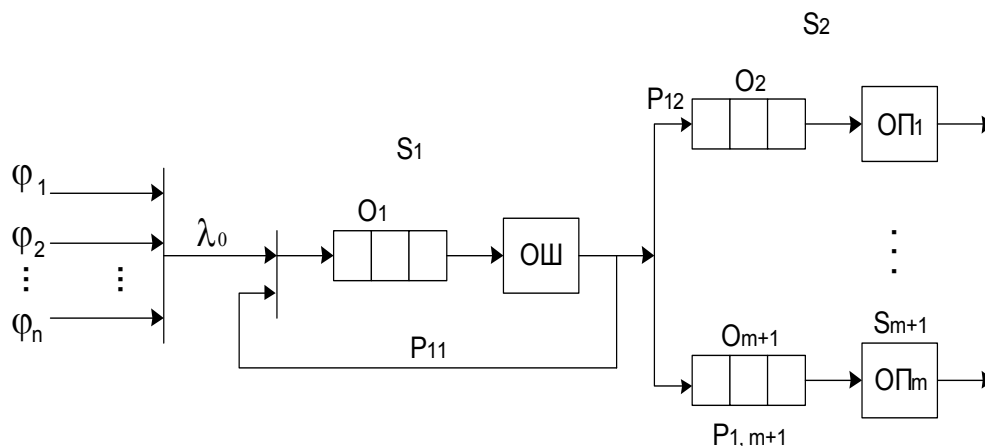


Рис.6.10. Разомкнутая сетевая модель системы с разделяемой памятью.

Рассмотренная модель предназначена для исследования многопроцессорных систем с интерфейсом, использующим принцип обмена с буферизацией сообщений, т.е. структуры, изображённой на рис.6.9. СМО S_1 моделирует ОШ, а S_2, \dots, S_{m+1} - модули общей памяти. В процессе обращения микропроцессоров к модулям общей памяти каждая заявка проходит два этапа. Вначале она требует доступа к ОШ, затем к одному из регистров буферного накопителя размещённого в коммутаторах КОП модуля памяти. Буферный накопитель моделируется очередью O_j ($j=2, \dots, m+1$).

Если в момент выдачи заявки ОШ оказывается занятой, то процессор, её выдавший, переходит в режим ожидания, который длится до тех пор, пока не освободится ОШ. Режим ожидания моделируется очередью O_1 , которая считается неограниченной. Граф передач сети представляется в виде, изображённом на рис.6.11.

Рассмотрим методику определения влияния конфликтных

ситуаций, возникающих в ОШ и общей памяти, на время выполнения команды процессора. Очевидно, что время выполнения команды зависит от времени обращения к памяти. Поскольку процесс обращения к памяти состоит из двух составляющих – обслуживания в общей шине и обслуживания в общей памяти, то время доступа к памяти будет определяться временем пребывания заявки в СМО S_I и в СМО S_j , т.е. временем пребывания заявки в сети.

Интенсивность потока заявок на входе сети $\lambda_o = \sum_{i=1}^n \varphi_i$, где φ_i – интенсивность потока команд i -го процессора. Она может быть определена как $\varphi_i = N_i / T_i$, где N_i – среднее число команд обращения i -го микропроцессора к памяти за время решения задачи T_i .

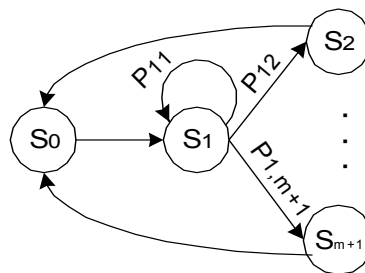


Рис.6.11. Граф передач сети, изображенной на рис.6.10.

Интенсивность потоков на входах СМО сети составит $\lambda_j = P_{1j} \cdot \lambda_1$, а поскольку $\lambda_1 = \lambda_o + P_{11}\lambda_1$, то отсюда следует: $\lambda_1 = \frac{\lambda_o}{1 - P_{11}}$.

Вероятности обращения заявок в СМО $S_j (j=2, \dots, m+1)$ могут быть определены как $P_{1j} = N_{1j} / N$, где N_{1j} – среднее число команд обращения в j -й модуль памяти за время решения задачи; N – суммарное число команд обращения к памяти микропроцессоров, т.е. $N = \sum_{j=2}^{m+1} N_j$.

В случае неограниченных длин очередей в СМО сети отказы в обслуживании будут отсутствовать, т.е. $P_{11}=0$, $\lambda_1=\lambda_o$.

Время пребывания заявки в СМО S_1 составит $u_1^o = \frac{\rho_1}{\lambda_1(1-\rho_1)}$,

где $\rho_1 = \lambda_1 t_B$ - нагрузка общей шины, t_B - время обмена данными по общей шине. Время пребывания заявки в СМО $S_j (j=2, \dots, m+1)$ составит

$u_j = \frac{\rho_j}{\lambda_j(1-\rho_j)}$, где $\rho_j = \lambda_j t_C$ - нагрузка модуля памяти; t_C - цикл памяти.

В случае ограничения длины очереди перед СМО S_j время пребывания определится как

$$u_j = \frac{1-\rho_j}{\lambda_j(1-\rho_j^{R+1})} \sum_{\tau=1}^R \tau \rho_j^\tau,$$

где R - суммарное число мест в очереди и в обслуживающем приборе.

В этом случае возможны отказы в обслуживании СМО S_j , причём вероятность отказа p_{11} будет складываться из вероятностей отказа в каждом из СМО S_2, \dots, S_{m+1} . Учитывая, что отказы начинаются с момента, когда $r=R$, получим:

$$P_{11} = \sum_{j=2}^{m+1} \frac{(1-\rho_j)\rho_j^R}{1-\rho_j^{R+1}}.$$

При обслуживании СМО S_1 с относительными фиксированными приоритетами время пребывания заявки, имеющий приоритет, в этой СМО определится через время задержки ω_f и время обслуживания обмена t_B :

$$u_1^f = \frac{\omega_o + \sum_{i=f+1}^n \rho_i \cdot \omega_i}{1 + \sum_{i=f}^n \rho_i} + t_B,$$

где $f=1, 2, \dots, n$ - классы приоритетов заявок, присвоенных процессорам. Время пребывания заявки в сети $U = \alpha_1 u_1 + \alpha_j j u_j$, где $\alpha_1 = \lambda_1 / \lambda_0$ и $\alpha_j = \lambda_j / \lambda_0$ - коэффициенты передачи сети, определяющие, сколько раз заявка,

поступившая в сеть, получает обслуживание в соответствующем узле.

Полученное время пребывания является временем обращения микропроцессора к памяти. Если в каждой из СМО сети длина очереди неограниченна, то это время составит

$$t_{MG} = \frac{\alpha_1 \rho_1}{\lambda_1(1 - \rho_1)} + \frac{\alpha_j \rho_j}{1 - \rho_j}.$$

Аналогично можно получить выражения t_{MG} для тех случаев, когда в узлах сети используются другие дисциплины обслуживания.

Если используется способ связи без буферизации сообщений, то общей шиной и, следовательно, общей памятью будет монопольно владеть только один вычислительный модуль, причём до тех пор, пока полностью не завершится обмен между ВМ и модулем памяти. Модель системы в этом случае преобразуется в одноузловую (рис.6.12). Время обращения к общей памяти

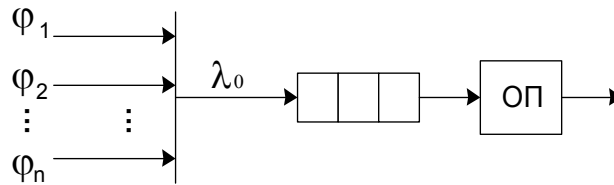


Рис.6.12. Модель системы с общей памятью и шинным интерфейсом без буферизации сообщений

$$t_{MG} = \frac{\rho}{\lambda_o(1 - \rho)}, \text{ где } \lambda_o = \sum_{i=1}^n \varphi_i; \quad \rho = \lambda_o \cdot t_B.$$

6.4.2. Микроанализ конфликтов в системах с общей шиной и распределённой памятью

Построим модель такой системы в предположении, что она состоит из n процессоров, каждый из которых имеет локальную память. Часть ЛП,

доступная всем микропроцессорам системы, определена как общая.

В процессе вычислительный модуль вначале запрашивает ОШ. Получив доступ к шине, ВМ-источник выставляет адреса, управляющие сигналы и данные для их записи (считывания) в область общей памяти ВМ-приёмника. Будем считать, что обращение к общей памяти производится прямым доступом, т.е. ВМ-источник обладает приоритетом по отношению ВМ-приёмнику.

Для передачи одного слова данных от одного вычислительного модуля другому требуется один цикл шины при межмодульной связи без буферизации сообщений. При связи с буферизацией для записи требуется один цикл шины, для чтения – два цикла. Средняя длительность цикла шины в этом случае определяется выражением (6.27). Микропроцессоры, потребовавшие доступа к занятой ОШ, переходят в режим ожидания.

Вначале рассмотрим модель системы, в которой используется принцип связи по ОШ без буферизации сообщений. На первой фазе обслуживания (рис.6.13) она содержит модель ОШ (СМО S_1), а на второй – модель модуля памяти (СМО S_2, \dots, S_{n+1}). Число СМО, моделирующих память, равно числу ВМ, т.к. вся память системы как локальная, так и общая распределена по вычислительным модулям.

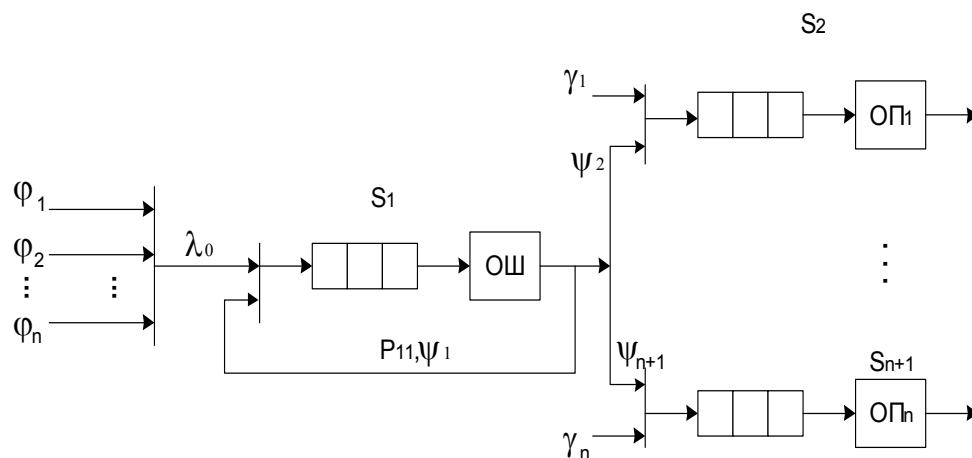


Рис.6.13. Модель системы с распределённой памятью.

Будем считать, что заявка, поступившая из потока ψ_1, \dots, ψ_n , принимается на обслуживание немедленно (с относительным приоритетом). Если СМО $S_j (j=2, \dots, n+1)$ занята обслуживанием ранее поступившей заявки из этого же потока, то она получает отказ в обслуживании и возвращается в конец очереди к СМО S_1 . Таким образом, моделируется процесс прямого доступа к памяти со стороны внешних, по отношению к рассматриваемому, вычислительных модулей. Суммарный поток отказов узлов S_2, \dots, S_{n+1} составит вероятность отказа p_{11} , причём $P_{11} = \sum_{j=2}^{n+1} P_{omkj}$, где P_{omkj} - вероятность отказа СМО S_j . Поскольку отказы в СМО S_j возникают в том случае, когда обслуживающий прибор (в данном случае модуль памяти) занят обслуживанием ранее поступившей заявки из потока ψ_j , то P_{omkj} определяется вероятностью занятия прибора заявками из этого потока. Вероятность занятия одноканального прибора равна ρ_j , поэтому

$$P_{omkj} = \rho_j, \text{ а } P_{11} = \sum_{j=2}^{n+1} P_j = \sum_{j=2}^n \psi_j t_c.$$

Интенсивность заявок на входе СМО

$$\lambda_1 = \frac{\lambda_o}{1 - \rho_{11}}, \text{ где } \lambda_o = \sum_{i=1}^n \phi_i.$$

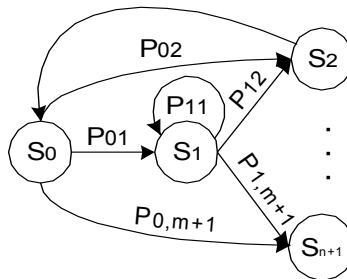


Рис.6.14. Граф передач сети, изображённой на рис.6.13.

Заявка, обслуженная СМО S_1 , с вероятностью p_{1j} поступает на обслуживание в один из узлов S_2, \dots, S_{n+1} , причём, как и в предыдущих

моделях, $p_{lj}=N_j/N$. Интенсивность потока заявок во второй фазе обслуживания $\psi_j=p_{lj}\cdot\lambda_l$.

Заявки из потока λ_l проходят две фазы обслуживания: вначале в СМО S_l , моделирующем ОШ, затем в одном из узлов S_2, \dots, S_{n+1} .

Время пребывания в СМО S_l при беспriorитетной дисциплине обслуживания

$$u_1 = \frac{\rho_1}{\lambda_1(1-\rho_1)},$$

где $\rho_l=\lambda_l \cdot t_B$ - загрузка ОШ.

Во второй фазе обслуживания (СМО S_2, \dots, S_{n+1}) заявки из потока φ_i , а следовательно, и заявки из потока φ_j получают приоритетное обслуживание. Поскольку приоритет у всех заявок из потока φ_j одинаков ($f=1$), то в соответствии с выражением (6.24) получим время ожидания их обслуживания

$$\omega_j = \frac{\omega_0}{1-\rho_j}.$$

Тогда время пребывания заявки в СМО $S_j(j=2, \dots, n+1)$ определится как $u_j = \omega_j + \rho_j$, где $\rho_j=\psi_j \cdot t_C$ - загрузка памяти заявками из потока φ_j . Следовательно,

$$u_j = \frac{\omega_0}{1-\rho_j} + \rho_j,$$

где ω_0 в выражениях является остаточным временем обслуживания заявок из потока λ_i . Оно определяется как доля времени, в течение которого прибор будет занят заявками из этого потока, т.е. $\omega_0=\rho_j$, следовательно,

$$u_j = \frac{\rho_j}{1-\rho_j} + \rho_j.$$

Поток команд обращения i -го микропроцессора к локальной памяти

моделируется потоком заявок γ_i .

Заявка, заставшая СМО $S_i(i=2, \dots, n+1)$ занятым, становится в одноступенчатую очередь и ожидает до тех пор, пока не будут обслужены заявки из потока ψ_i , поскольку они обладают более высоким приоритетом. Местом в очереди является непосредственно процессор (один из его регистров), выдавший заявку на обращение в локальную память, т.к. он переходит в режим ожидания, который будет длиться до тех пор, пока память не освободится.

Таким образом, моделью процесса обмена с локальной памятью является СМО с ограниченной длиной очереди ($R=2$). В соответствии с (6.21) время пребывания заявки и, следовательно, время обращения к локальной памяти t_{ML} составит:

$$t_{ML} = \frac{(1-\rho_i)\rho_i}{\gamma_i(1-\rho_i^3)} + \frac{(1-\rho_j)\rho_i^2}{\gamma_i(1-\rho_i^3)}.$$

Преобразив, получим

$$t_{ML} = \frac{\rho_i(1-\rho_i)\rho_i^2}{\gamma_i(1-\rho_i^3)},$$

где $\rho_i = \gamma_i t_C$; $\gamma_i = N_{MLi}/N_i$.

Здесь N_{MLi} - число команд обращения микропроцессора в локальную память; N_i - общее число команд в решаемой задаче i -м вычислительным модулем.

Общее время пребывания заявки из потока ϕ_i в сети определится как $U = \alpha_i u_i + \alpha_j u_j$, следовательно, время обращения в общую память составит

$$t_{MG} = \frac{\alpha_i \rho_i}{\lambda_i(1-\rho_i)} + \frac{\alpha_j \rho_j}{1-\rho_j} + \alpha_j \rho_j.$$

Системы с распределённой памятью, в межмодульном интерфейсе которых использован принцип связи с буферизацией сообщений, отличается наличием буферного накопителя в каждом модуле памяти

ёмкостью R . Модель будет отличаться от предыдущей наличием ограниченной очереди для заявок из потока ψ_j перед СМО $S_j (j=2, \dots, n+1)$.

При исследовании системы с помощью предложенной модели следует учесть, что дисциплина обслуживания перед СМО S_j может быть либо *FIFO*, либо приоритетная. Во втором случае более высокий приоритет следует давать заявкам из потока $\psi_j (j=2, \dots, n+1)$, чтобы не перегружать ОШ повторными передачами сообщений, получивших отказ в обслуживании СМО S_2, \dots, S_{n+1} из-за ограниченной длины очереди. Чаще всего число мест в очередях как для потока γ_j , так и для потока ψ_j , ограничивают до одного.

В заключение следует отметить, что построение модели вычислительной системы является достаточно трудоёмким процессом, т.к. требует тщательного изучения вычислительных процессов. Однако на этом создание модели не заканчивается. Возникают вопросы – достоверна ли модель; насколько хорошо она представляет моделируемую систему. Наилучший ответ на эти вопросы даст реальная вычислительная система. Однако такой ответ можно получить только через годы, которые уйдут на её изготовление. Самый простой метод проверки достоверности заключается в построении некоторой другой модели, например, марковский, результаты которой не должны противоречить полученным результатам.

Другой вопрос связан с точностью модели и её калибровкой. Точность оценивается разницей между реальными характеристиками системы (например, реальной производительностью) и характеристиками, полученными моделью. Чем меньше ошибка, тем точнее модель. Однако, как и в первом случае, на проверку точности могут уйти годы. Наиболее приемлемым методом проверки точности является построение другой модели; более детально описывающей вычислительную систему. В

качестве такой модели может выступать, например, имитационная.

Если точность модели неудовлетворительна, то модель должна быть изменена, а процесс проверки повторён. Эта операция называется калибровкой. Процедура калибровки оканчивается, как только исследователь найдёт точность модели удовлетворительной. Только после этого она может быть рекомендована для проведения исследований класса вычислительных систем, на которые модель направлена.

Литература

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Математический энциклопедический словарь / Гл. ред. Ю. В. Прохоров. – М. : Сов. Энциклопедия, 1988. – 847 с.
2. *Калашников В. В.* Математические методы построения стохастических моделей обслуживания / В. В. Калашников, С. Т. Рачев С.Т. – М. : Наука. Гл. ред. физ.-мат. лит. 1988. – 312 с.
3. *Таха Х.* Введение в исследование операций: В 2-х книгах. Кн.2 / Х. Таха. – М.: Мир, 1985. – 496 с.
4. *Клейнрок Л.* Теория массового обслуживания / Л. Клейнрок. – М. : Машиностроение, 1979. – 432 с.
5. Ивченко Г. И. Теория массового обслуживания / Г. И. Ивченко, В. А. Каштанов, И. Н. Коваленко. – М. : Высш. шк., 1982. – 256 с.
6. Матвеев В.Ф. Системы массового обслуживания / В. Ф. Матвеев, В. Г. Ушаков. – М.: Изд-во Моск. ун-та, 1984. – 238 с.
7. Математика для экономистов: В 6 т./ Под ред. А. Ф.Тарасюка. — М. : ИНФРА-М, 2000. — (серия "Высшее образование"). Т.6: Чернов В. П., Ивановский В. Б. Теория массового обслуживания. – 158 с.
8. *Вентцель Е. С.* Теория случайных процессов и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. – М. : Высш. шк., 2000. – 383 с.
9. *Вентцель Е. С.* Теория вероятностей: Учеб. для вузов. 6-е изд. стер. / Е. С. Вентцель, – М. : Высш. шк., 1999. – 576 с.
10. *Кремер Н. Ш.* Теория вероятностей и математическая статистика: Учебник для вузов / Н. Ш. Кремер. – М. : ЮНИТИ-ДАНА, 2000. – 543 с.

11. *Колемаев В. А.* Теория вероятностей и математическая статистика: Учебник / В. А. Колемаев, В. Н. Калинина. Под ред. В. А. Колемаева. – М. : ИНФРА-М, 1997. – 302 с.
12. *Матвеев Н. М.* Сборник задач и упражнений по обыкновенным дифференциальным уравнениям / Н. М. Матвеев. – Минск: "Вышэйш. школа", 1967. – 308 с.
13. *Тараканов К. В.* Аналитические методы исследования систем / К. В. Тараканов, Л. А. Овчаров, А. Н. Тырышкин. – М.: Сов. радио, 1974. – 240 с.
14. *Марков А. А.* Моделирование информационно-вычислительных процессов: Учебное пособие для вузов / А. А. Марков. — М. : Изд-во МГТУ им. Н. Э. Баумана, 1999. – 360 с.
15. *Феллер В.* Введение в теорию вероятностей и ее приложения. В 2-х томах. Т.1: Пер с англ. / В. Феллер. – М. : Мир, 1984. – 528 с.
16. *Дегтерев Ю. И.* Системный анализ и исследование операций / Ю. И. Дегтярев. – М. : Высш. шк., 1996. – 335 с.
17. *Корн Г.* Справочник по математике. Для научных работников и инженеров / Г. Корн, Т. Корн. – М.: Наука, 1978. – 832 с.
18. *Волков И. К.* Случайные процессы: Учеб. для вузов / И. К. Волков, С. М. Зуев, Г. М. Цветкова. Под ред. В. С. Зарубина, А. П. Крищенко. – М. : Изд-во МГТУ им. Н.Э.Баумана, 2000. – 448 с.
19. *Лабскер Л. Г.* Вероятностное моделирование в финансово-экономической области / Л.Г. Лабскер. – М. : Альпина Паблишер, 2002. – 224 с.
20. *Мартынов Н.Н.* MATLAB 5.x. Вычисления, визуализация, программирование / Н. Н. Мартынов, А. П. Иванов. – М. : КУДИЦ-ОБРАЗ, 2000. – 336 с.
21. *Говорухин В.* Компьютер в математическом исследовании. Учебный курс / В. Говорухин, В. Цибулин. — СПб. : Питер, 2001. – 624 с.
22. *Дьяконов В. П.* MATLAB 6/6.1/6.5 +Simulink 4/5. Основы применения. Полное руководство пользователя / В. П. Дьяконов. – М. : СОЛОН-Пресс. — 2002. — 768 с.

23. Кетков Ю. Л. MATLAB 7: программирование, численные методы / Ю. Л. Кетков, А. Ю. Кетков, М. М. Шульц. – СПб. : БХВ-Петербург, 2005. – 752 с.
24. Клейнрок Л. Вычислительные системы с очередями. – М.: Мир, 1979.-600с.
25. Основы теории вычислительных систем/ Под ред.С.А.Майорова. – М.: Высшая школа, 1978.-408с.
26. Феррари Д. Оценка производительности вычислительных систем. – М.: Мир, 1981. –573с.
27. Байцер Б. Микроанализ производительности вычислительных систем. –М.: Радио и связь, 1983.-360с.
28. Компьютеры на СБИС: В 2-х кн. Кн. 2: Пер. с япон./ Мотоока Т., Сакаuti М.,Харикоси Х. И др.- М.: Мир, 1988.- 336с.
29. Сетевые операционные системы / В.Г. Олифер, Н.А. Олифер.- СПб.: Питер, 2001.- 544с.
30. Шагурин И.И., Бердышев Е.М. Процессоры семейства Intel P6. Архитектура, программирование, интерфейс.- М.: Горячая линия - Телеком, 2000.- 248 с.
31. Корнеев В.В.,Кисилёв А.В. Современные микропроцессоры.- М.: Нолидж, 1988.- 240 с.
32. Бикташев Р.А. Моделирование мультимикропроцессорных вычислительных систем: Учебное пособие. – Пенза: Пенз. политехн. ин-т, 1987. – 44 с.
33. Князьков В. С., Бикташев Р.А. Архитектура параллельных вычислительных систем.-Пенза: 1993.- 166с.