Раздел 5. ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ (МОДЕЛИ СЛУЧАЙНЫХ ПРОЦЕССОВ)

«В задаче из N уравнений всегда будет N+1 неизвестная» ***(Уравнения Снэйфу)***

При изучении сложных систем со стохастическим характером функционирования полезной математической моделью является *случайный процесс*, который развивается в зависимости от ряда случайных факторов. Примерами случайных процессов могут служить процессы поступления и передачи данных в телекоммуникационной сети, процессы выполнения задач и обмена данными с внешними устройствами в вычислительной системе и т.п.

Большинство моделей дискретных систем со стохастическим характером функционирования строится на основе моделей массового обслуживания, процессы в которых являются случайным и, во многих случаях, *марковскими* или некоторым образом связанные с марковскими процесссами. Поэтому для решения таких задач теории массового обслуживания может использоваться математический аппарат **теории марковских процессов**. Применение марковских процессов оказывается особенно эффективным и результативным *при исследовании систем и сетей массового обслуживания с накопителями ограниченной ёмкости*.

Математическое описание марковских процессов обычно представляется в виде систем дифференциальных (в случае нестационарного режима) или алгебраических (для стационарного режима) уравнений, решение которых, в общем случае, получить в явном виде не удается. Это обусловливает необходимость применения *численных методов* решения систем дифференциальных или алгебраических уравнений.

5.1. Понятие случайного процесса

Основными для случайных процессов являются понятия *состояния* и *перехода* из одного состояния в другое.

Случайный процесс находится в некотором ***состоянии***, если он полностью описывается значениями переменных, которые задают это состояние.

Процесс совершает ***переход*** из одного состояния в другое, если описывающие ее переменные изменяются от значений, задающих одно состояние, на значения, которые определяют другое состояние.

Случайный процесс состоит в том, что с течением времени процесс переходит из одного состояния в другое *заранее не известное состояние*.

Понятия «состояние» и «переход» используются как для описания случайного процесса, так и системы, в которой этот процесс протекает. Поэтому при моделировании реальных систем часто говорят о состоянии системы и переходе системы из одного состояния в другое.

Если множество состояний, в которых может находиться процесс *счётное*, то есть все возможные состояния могут быть пронумерованы, то соответствующий процесс называется ***случайным процессом с дискретными состояниями*** или просто ***дискретным случайным процессом***. В этом случае переменные, описывающие состояния случайного процесса, принимают либо целочисленные значения, либо вполне конкретные отделённые друг от друга дискретные значения. Обычно состояния дискретного случайного процесса определяются таким образом, чтобы каждое возможное состояние могло быть обозначено порядковым номером, при этом число возможных состояний системы может быть *конечным*: $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_n$ или *бесконечным*: $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_n, ...$ (иногда состояния нумеруются, начиная с нуля: $\mathbf{E}_0, \mathbf{E}_1, ..., \mathbf{E}_n, ...$). Для случайного процесса с дискретными состояниями характерен скачкообразный переход из одного состояния в другое (рис.5.1,а). Например, случайный процесс, протекающий в простейшей СМО с однородным потоком заявок, может быть представлен количеством заявок, находящихся в системе в произвольный момент времени. Тогда состояние $\mathbf{E}_k$ случайного процесса и, следовательно, самой системы будет означать, что в СМО находится ровно $k = 0, 1, 2, ...$ заявок.

Если множество состояний не может быть пронумеровано, то имеем ***случайный процесс с непрерывными состояниями*** или просто ***непрерывный случайный процесс***, для которого характерен плавный переход из состояния в состояние и который задаётся в виде непрерывной ункции времени: $E(t)$ (рис.5.1,б). Например, процесс изменения температуры некоторого объекта может рассматриваться как случайный процесс с непрерывными состояниями.

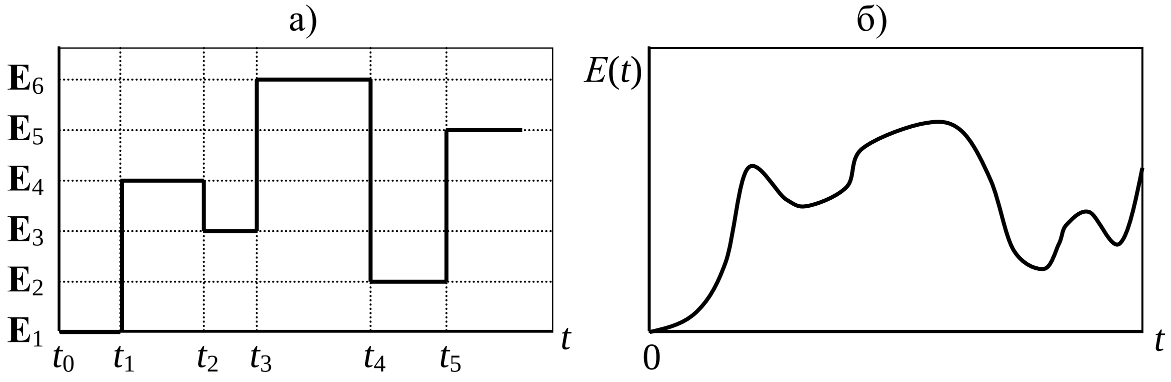


Рис.5.1. Процессы с дискретными (а) и непрерывными (б) состояниями

Поскольку модели массового обслуживания относятся к классу дискретных систем, то в дальнейшем будут рассматриваться только случайные процессы с дискретными состояниями.

При описании дискретных систем в терминах случайных процессов одним из основных этапов является этап *кодирования состояний*, заключающийся в определении состава переменных и их значений, используемых для описания состояний. Состав переменных в значительной мере определяется назначением разрабатываемой модели, зависящим от целей исследований.

5.1.1. Случайные процессы с дискретными состояниями

Предположим, что система может находиться в одном из состояний **E**$_1$, **E**$_2$, ... (часто состояния обозначаются просто номерами 1, 2,...). Пусть состояние системы меняется скачкообразно в зависимости от некоторого параметра t, причем переход из состояния в состояние является случайным. Будем называть параметр t – **временем** и считать, что t пробегает либо целые, либо действительные числа. Обозначим через $Z(t)$ случайный процесс, описывающий состояние системы в момент времени t.

Случайный процесс $Z(t)$ называется ***случайным процессом с дискретным временем***, если переходы из состояния в состояние возможны только в строго *определенные заранее фиксированные моменты времени*, которые можно пронумеровать: $t_1, t_2, \dots$.

Если промежуток времени между переходами из состояния в состояние является *случайным* и переход возможен в любой заранее не известный момент времени t, то процесс называется ***случайным процессом с непрерывным временем***.

Процесс с дискретным временем имеет место либо когда структура системы такова, что ее состояния могут изменяться только в заранее определенные моменты времени, либо когда предполагается, что для описания процесса достаточно знать состояние системы в отдельные моменты времени. Тогда эти моменты можно пронумеровать и говорить о состоянии **E**$_i$ в момент t_k или просто в момент k $(k = 0, 1, 2, \dots)$.

Процессы с дискретным временем называются **стохастическими последовательностями** или **случайными цепями**.

Случайные процессы с дискретными состояниями могут изображаться в виде **графа переходов** (состояний), в котором вершины соответствуют состояниям, а ориентированные дуги – переходам из одного состояния в другое.

Граф переходов называется ***размеченным***, если на дугах графа указаны условия перехода в виде *вероятностей переходов* (для процессов с дискретным временем) или *интенсивностей переходов* (для процессов с непрерывным временем).

Состояния **E**$_i$ могут быть:

- **невозвратными**, если процесс после какого-то числа переходов непременно покидает их;
- **поглощающими**, если случайный процесс, достигнув этих состояний прекращается.

Случайный процесс называется ***транзитивным***, если из любого состояния можно перейти за то или иное число шагов в любое другое состояние и вернуться в исходное.

5.1.2. Понятие марковского случайного процесса

Случайный процесс называется ***марковским***, если вероятность любого состояния в будущем зависит только от его состояния в настоящем и не зависит от того, когда и каким образом процесс оказался в этом состоянии.

Описывающий поведение системы процесс $Z(t)$ называется **цепью Маркова**.

Для того чтобы случайный процесс *с непрерывным временем* был *марковским*, необходимо, чтобы интервалы времени между соседними переходами из состояния в состояние были распределены по *экспоненциальному закону*. Для доказательства последнего утверждения воспользуемся следующими рассуждениями.

Пусть время нахождения случайного процесса в некотором состоянии $\mathbf{E}_i$ до его перехода в другое состояние $\mathbf{E}_j$ распределено по экспоненциальному закону с функцией распределения $F_{ij}(\tau) = 1 - e^{-\alpha_{ij}\tau}$, где α_{ij} - параметр распределения, характеризующий частоту перехода из состояния $\mathbf{E}_i$ в состояние $\mathbf{E}_j$ и определяемый как величина, обратная среднему времени нахождения случайного процесса в состоянии $\mathbf{E}_i$ до момента его перехода в состояние $\mathbf{E}_j$. Вычислим вероятность того, что случайный процесс перейдет в состояние $\mathbf{E}_j$ в течение интервала времени $\Delta\tau$ при условии, что в состоянии $\mathbf{E}_i$ процесс уже находится в течение времени τ_0. Эта условная вероятность равна

$$P_{ij}(\Delta\tau \mid \tau \ge \tau_0) = \Pr(\tau_0 \le \tau \le \tau_0 + \Delta\tau \mid \tau \ge \tau_0) =$$

$$= \frac{\Pr(\tau_0 \le \tau \le \tau_0 + \Delta\tau)}{\Pr(\tau \ge \tau_0)} = \frac{F(\tau_0 + \Delta\tau) - F(\tau_0)}{1 - F(\tau_0)} = 1 - e^{-\alpha_{ij}\Delta\tau}.$$

Из последнего выражения следует, что вероятность перехода из одного состояния в другое зависит только от исходного состояния $\mathbf{E}_i$ и не зависит от интервала времени τ_0, то есть от того, как долго находился процесс в состоянии $\mathbf{E}_i$, а также от того, какие состояния предшествовали состоянию $\mathbf{E}_i$. Другими словами, поведение случайного процесса не зависит от предыстории и определяется только его состоянием в настоящий момент, то есть процесс является марковским.

Еще одно замечательное **свойство экспоненциального распределения** вытекает из полученного выражения, а именно: если время нахождения случайного процесса в некотором состоянии $\mathbf{E}_i$ до его перехода в другое состояние $\mathbf{E}_j$ распределено по экспоненциальному закону с параметром α_{ij}, то *интервал времени от любого случайного момента времени до момента перехода* в состояние $\mathbf{E}_j$ имеет *такое же экспоненциальное распределение с тем же параметром* α_{ij}. Эта особенность является следствием отсутствия последействия, присущего всем процессам с экспоненциальным распределением времени нахождения в том или ином состоянии.

Таким образом, безусловная $P_{ij}(\Delta\tau)$ и условная $P_{ij}(\Delta\tau \mid \tau \ge \tau_0)$

вероятности перехода в другое состояние за время $\Delta\tau$ для марковского процесса одинаковы и равны

$$P_{ij}(\Delta\tau) = P_{ij}(\Delta\tau \mid \tau \ge \tau_0) = 1 - e^{-\alpha_{ij}\Delta\tau}.$$

Пусть интервал времени $\Delta\tau$ достаточно мал. Тогда, разлагая $e^{-\alpha_{ij}\Delta\tau}$ в ряд по степеням $\alpha_{ij}\Delta\tau$ при $\Delta\tau \to 0$ и пренебрегая величинами высшего порядка малости, получим вероятность перехода из одного состояния в другое за бесконечно малый интервал времени:

$$P_{ij}(\Delta\tau) = 1 - (1 - \alpha_{ij}\Delta\tau) = \alpha_{ij}\Delta\tau. \qquad (5.1)$$

5.2. Параметры и характеристики марковского случайного процесса

5.2.1. Параметры марковского случайного процесса

Для описания марковского случайного процесса с дискретными состояниями используется следующая совокупность параметров:

- **перечень состояний** $\mathbf{E}_1,\ldots,\mathbf{E}_n$, в которых может находиться случайный процесс;
- **матрица переходов**, описывающая переходы случайного процесса между состояниями в виде:
  - *матрицы вероятностей переходов* **Q** для процессов *с дискретным временем*;
  - *матрицы интенсивностей переходов* **G** для процессов *с непрерывным временем*;
- **начальные вероятности** $p_1(0),\ldots, p_n(0)$.

Для **определения перечня состояний** случайного процесса необходимо корректно решить задачу кодирования состояний, которое зависит от смысла, вкладываемого в понятие «состояние» для каждой конкретной системы. Так, например, состояние некоторой системы массового обслуживания (а, следовательно, и случайного процесса, протекающего в ней) может быть задано числом заявок, находящихся в системе в данный момент времени, а состояние сети массового обслуживания – распределением числа заявок по всем узлам сети.

Для случайных процессов с дискретным временем изменения состояний происходят только в определенные моменты времени $t_1, t_2, \ldots, t_k, \ldots$. Переходы между состояниями описываются ***вероятностями переходов***. Если непосредственный переход из одного состояния в другое невозможен, то вероятность, соответствующая данному переходу, равна нулю. Обозначим через q_{ij} условную вероятность того, что в момент времени t_{k+1} случайный процесс перейдет в состояние $\mathbf{E}_j$ при условии, что в момент t_k процесс находился в состоянии $\mathbf{E}_i$. Если переход из состояния

$\mathbf{E}_i$ в $\mathbf{E}_j$ зависит только от этих двух состояний, то есть условная вероятность q_{ij} не изменяется при дополнительной информации о поведении процесса до момента t_k, получим цепь Маркова.

Цепь Маркова называется ***однородной***, если вероятности переходов не зависят от момента времени t_k, и ***неоднородной***, если вероятности переходов являются функциями t_k, то есть $q_{ij} = q_{ij}(k)$.

Вероятности переходов задаются в виде квадратной ***матрицы вероятностей переходов*** $\mathbf{Q} = [q_{ij} \mid i, j = \overline{1, n}]$, элементы которой удовлетворяют условиям:

$$0 \le q_{ij} \le 1; \quad \sum_{j=1}^{n} q_{ij} = 1 \qquad (i, j = \overline{1, n}). \tag{5.2}$$

Матрица, элементы которой удовлетворяют указанным условиям, называется ***стохастической***.

Последнее условие в виде суммы элементов каждой строки матрицы вероятностей переходов, равной единице, означает, что в момент времени t_k случайный процесс с вероятностью единица выполнит переход в одно из n возможных состояний, включая то же самое состояние, из которого этот переход осуществляется, то есть процесс может остаться в том же состоянии.

Для случайных процессов с непрерывным временем время между переходами из одного состояния в другое случайно. Это означает, что вероятность перехода из одного состояния в другое не может быть задана, поскольку вероятность такого перехода точно в произвольный момент времени t равна нулю. Для описания переходов между состояниями случайного процесса с непрерывным временем вместо вероятностей переходов вводится параметр, называемый *интенсивностью перехода*.

Интенсивность перехода g_{ij} из состояния $\mathbf{E}_i$ в состояние $\mathbf{E}_j$ определяется как предел отношения вероятности перехода $P_{ij}(\Delta\tau)$ системы за промежуток времени $\Delta\tau$ из $\mathbf{E}_i$ в $\mathbf{E}_j$ к длине этого промежутка:

$$g_{ij} = \lim_{\Delta\tau \to 0} \frac{P_{ij}(\Delta\tau)}{\Delta\tau} \qquad (i, j = \overline{1, n};\ i \ne j). \tag{5.3}$$

Отсюда следует, что вероятность перехода за бесконечно малый промежуток времени $\Delta\tau$ равна: $g_{ij}\Delta\tau \quad (i \ne j)$. Вероятность двух и более переходов за время $\Delta\tau$ имеет порядок $(\Delta\tau)^2$ и выше и предполагается бесконечно малой величиной.

Если интенсивности переходов постоянны и не зависят от времени t, то есть от того, в какой момент начинается промежуток $\Delta\tau$, то марковский процесс называется ***однородным***. Если интенсивности g_{ij} представляют собой функции времени t, процесс называется ***неоднородным***.

В дальнейшем будем рассматривать только однородные марковские процессы.

Интенсивности переходов задаются в виде квадратной матрицы $\mathbf{G} = [g_{ij} \mid i, j = \overline{1,n}]$, называемой ***матрицей интенсивностей переходов***, диагональные элементы которой определяются из условия:

$$\sum_{j=1}^{n} g_{ij} = 0 \quad (i = \overline{1,n}),$$

откуда

$$g_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^{n} g_{ij} \quad (i, j = \overline{1,n}). \tag{5.4}$$

Матрица, в которой сумма элементов в каждой строке равна нулю, называется ***дифференциальной***.

Выше было показано, что в случае экспоненциального закона распределения времени нахождения случайного процесса в некотором состоянии вероятность перехода из одного состояния в другое за бесконечно малый интервал времени определяется выражением (5.1) и равно $P_{ij}(\Delta\tau) = \alpha_{ij}\Delta\tau$. Отсюда следует, что *интенсивность перехода представляет собой параметр экспоненциального распределения*:

$$g_{ij} = \lim_{\Delta\tau \to 0} \frac{P_{ij}(\Delta\tau)}{\Delta\tau} = \alpha_{ij}.$$

Начальные вероятности $p_1(0), \ldots, p_n(0)$, где $p_i(0)$ – вероятность того, что в момент времени $t = 0$ система находится в состоянии $\mathbf{E}_i$ $(i = 1, \ldots n)$, задают состояние системы в начальный момент времени $t = 0$.

Начальные вероятности необходимы при изучении переходных процессов, когда исследуемая система работает в нестационарном режиме. Если марковский процесс обладает эргодическим свойством, что означает работу моделируемой системы в установившемся режиме, то, как будет показано ниже, стационарные характеристики (вероятности) не зависят от начальных вероятностей и, следовательно, могут быть не заданы.

5.2.2. Характеристики марковского случайного процесса

Изучение случайных процессов заключается в определении вероятностей того, что в момент времени t система находится в том или ином состоянии. Совокупность таких вероятностей, описывающих состояния системы в различные моменты времени, дают достаточно полную информацию о протекающем в системе случайном процессе.

Рассмотрим систему с конечным числом состояний: $\mathbf{E}_1$, …, $\mathbf{E}_n$. Обозначим через $p_i(t)$ вероятность того, что в момент времени t система находится в состоянии $\mathbf{E}_i$: $p_i(t) = \Pr\{Z(t) = E_i\}$.

В любой момент времени t система может находиться в одном из n

возможных состояний, то есть для любого момента времени t выполняется условие:

$$\sum_{i=1}^{n} p_i(t) = 1, \tag{5.5}$$

которое называется ***нормировочным***.

Совокупность вероятностей $p_i(t)$ может быть представлена вектором с числом координат, равным числу возможных состояний системы:

$$P(t) = \{p_1(t),\ldots,p_n(t)\},$$

причем

$$0 \le p_i(t) \le 1; \quad \sum_{i=1}^{n} p_i(t) = 1. \tag{5.6}$$

Вектор, обладающий свойствами (5.6), называется ***стохастическим***.

Стохастический вектор называется ***вектором состояний***, если его компоненты представляют собой вероятности состояний системы.

Вектор состояний $P(t) = \{p_1(t),\ldots,p_n(t)\}$ является основной характеристикой марковского случайного процесса. На основе полученных значений вероятностей состояний случайного процесса, протекающего в исследуемой системе, могут быть рассчитаны представляющие интерес реальные характеристики системы, например для системы массового обслуживания могут быть рассчитаны длины очередей заявок.

5.3. Методы расчета марковских моделей

5.3.1. Эргодическое свойство случайных процессов

Если по истечении достаточно большого промежутка времени вероятности состояний стремятся к предельным значениям $p_1,\ldots,p_n$, не зависящим от начальных вероятностей $p_1(0),\ldots,p_n(0)$ и от текущего момента времени t, то говорят, что случайный процесс обладает ***эргодическим свойством***. Таким образом, для процессов, обладающих эргодическим свойством:

$$\lim_{t\to\infty} P(t) = P(\infty) = \mathbf{P},$$

где $\mathbf{P} = (p_1,\ldots,p_n)$ – вектор вероятностей состояний системы, называемых ***стационарными вероятностями***.

В системе, описываемой марковским случайным процессом, обладающим эргодическим свойством, при $t \to \infty$ устанавливается некоторый предельный режим, при котором характеристики функционирования системы не зависят от времени. В этом случае говорят, что система работает в **установившемся** или **стационарном режиме**. Если характеристики функционирования системы зависят от времени, то имеем

неустановившийся режим.

Отметим, что для стационарных вероятностей p_i должно выполняться нормировочное условие (5.5).

При рассмотрении случайных процессов возникает вполне резонный вопрос: *когда случайный процесс обладает эргодическим свойством*?

Случайный процесс *с дискретным временем* обладает *эргодическим свойством*, если матрица вероятностей переходов **не является** *периодической* или *разложимой*.

Матрица является **разложимой**, если она может быть приведена к одному из следующих видов:

$$1)\ \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \qquad 2)\ \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \qquad 3)\ \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix},$$

где **A, B, C, D**– ненулевые квадратные подматрицы; **0** – нулевая квадратная подматрица.

В первом случае состояния, соответствующие подмножествам **A** и **D**, называются **замкнутыми**, так как система, находясь в каком-то состоянии одного из этих подмножеств, никогда не сможет перейти в какое-либо состояние другого подмножества. Состояния, соответствующие подмножеству **D** во втором случае и подмножеству **A** в третьем случае, называются **невозвратными**, поскольку после того, как процесс покинет эти состояния, невозможен обратный переход в эти состояния из состояний, соответствующих другим подмножествам.

Матрица является **периодической**, если она может быть приведена к виду:

$$\begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}.$$

Случайный процесс в этом случае будет по очереди переходить из состояний, соответствующих **B**, в состояния, соответствующие **C**.

Итак, если матрица вероятностей переходов $\mathbf{Q} = [q_{ij} \mid i, j = \overline{1,n}]$, случайного процесса с дискретным временем не является периодической или разложимой, то процесс обладает эргодическим свойством:

$$\lim_{k\to\infty} p_i(k) = p_i \qquad (i = \overline{1,n}). \tag{5.7}$$

Транзитивный случайный процесс с *непрерывным временем* и *конечным* числом состояний, среди которых нет невозвратных и поглощающих состояний, всегда обладает эргодическим свойством:

$$\lim_{t\to\infty} p_i(t) = p_i \qquad (i = \overline{1,n}). \tag{5.8}$$

5.3.2. Марковские процессы с дискретным временем

Для однородного марковского процесса с дискретным временем вероятности состояний на момент времени t_k определяются на основе следующего рекуррентного выражения:

$$p_j(k) = \sum_{i=1}^{n} p_i(k-1)\, q_{ij} \quad (j = \overline{1,n};\ k = 1,\ 2,\ldots). \tag{5.9}$$

Если рассматриваемый марковский процесс обладает эргодическим свойством, то, согласно (5.7), при $k \to \infty$ вероятности состояний $p_i(k)$ стремятся к стационарным значениям p_i, не зависящим от момента времени t_k и начальных вероятностей $p_i(0)$. С учётом этого, выражение (5.9) может быть преобразовано к виду:

$$p_j = \sum_{i=1}^{n} p_i q_{ij} \qquad (j = \overline{1,n}) \tag{5.10}$$

а нормировочное условие (5.5) примет вид:

$$\sum_{i=1}^{n} p_i = 1. \tag{5.11}$$

Уравнения (5.10) с условием (5.11) образуют систему линейных алгебраических уравнений для расчёта стационарных вероятностей состояний марковского процесса, которая обладает единственным решением, если **Q** – эргодическая матрица.

Доказательство выражения (5.9).

Рассмотрим однородный марковский процесс с дискретным временем, который может находиться в одном из n возможных состояний: **E**$_1$, …, **E**$_n$. Вероятности переходов q_{ij} заданы в виде матрицы переходов $\mathbf{Q} = [q_{ij} \mid i,\ j = \overline{1,n}]$, а начальные вероятности на момент времени $t_0 = 0$ в виде вектора $P = \{p_1(0),\ldots,\ p_n(0)\}$.

Найдем вероятности состояний марковского процесса после первого шага, то есть на момент времени t_1. По формуле полной вероятности получим:

$$\begin{cases} p_1(1) = p_1(0)q_{11} + p_2(0)q_{21} + \ldots + p_n(0)q_{n1}; \\ p_2(1) = p_1(0)q_{12} + p_2(0)q_{22} + \ldots + p_n(0)q_{n2}; \\ \qquad . \quad . \quad . \\ p_n(1) = p_1(0)q_{1n} + p_2(0)q_{2n} + \ldots + p_n(0)q_{nn}, \end{cases}$$

или в компактной форме:

$$p_j(1) = \sum_{i=1}^{n} p_i(0) q_{ij} \qquad (j = \overline{1,n}).$$

Вероятности состояний после второго шага на момент времени t_2 определяются аналогично:

$$p_j(2) = \sum_{i=1}^{n} p_i(1) q_{ij} \qquad (j = \overline{1,n}).$$

После k-го шага на момент времени t_k $(k = 1, 2, \ldots)$ вероятности состояний будут определяться как

$$p_j(k) = \sum_{i=1}^{n} p_i(k-1)\, q_{ij} \qquad (j = \overline{1, n}),$$

что и требовалось доказать.

Пример. Рассмотрим систему, которая состоит из двух устройств $У_1$ и $У_2$, каждое из которых может находиться в одном из двух состояний: **0** – выключено и **1** – включено. В определённые моменты времени устройства могут включаться или выключаться. Выделим возможные состояния системы:

$\mathbf{E}_i$	$\mathbf{E}_0$	$\mathbf{E}_1$	$\mathbf{E}_2$	$\mathbf{E}_3$
$У_1$	0	1	0	1
$У_2$	0	0	1	1

Состояние $\mathbf{E}_0$ соответствует простою системы, когда оба устройства выключены, а состояние $\mathbf{E}_3$ соответствует случаю, когда оба устройства включены.

Положим, что заданы вероятности переходов в виде матрицы

$\mathbf{Q}$ =	$\mathbf{E}_0$	$\mathbf{E}_1$	$\mathbf{E}_2$	$\mathbf{E}_3$
$\mathbf{E}_0$	0	0,2	0,5	0,3
$\mathbf{E}_1$	0,5	0	0,1	0,4
$\mathbf{E}_2$	0,5	0	0	0,5
$\mathbf{E}_3$	0	0,4	0,6	0

и начальные вероятности $p_0(0) = 0{,}8$; $p_1(0) = 0{,}2$; $p_2(0) = 0$; $p_3(0) = 0$.

Определим вероятности нахождения системы в том или ином состоянии на различные моменты времени.

Согласно выражению (5.9) вероятности состояний системы:

- на момент времени t_1:

$p_0(1) = p_0(0)q_{00} + p_1(0)q_{10} + p_2(0)q_{20} + p_3(0)q_{30} = 0{,}1;$

$p_1(1) = p_0(0)q_{01} + p_1(0)q_{11} + p_2(0)q_{21} + p_3(0)q_{31} = 0{,}16;$

$p_2(1) = p_0(0)q_{02} + p_1(0)q_{12} + p_2(0)q_{22} + p_3(0)q_{32} = 0{,}42;$

$p_3(1) = p_0(0)q_{03} + p_1(0)q_{13} + p_2(0)q_{23} + p_3(0)q_{33} = 0{,}32;$

- на момент времени t_2:

$p_0(2) = p_0(1)q_{00} + p_1(1)q_{10} + p_2(1)q_{20} + p_3(1)q_{30} = 0{,}29;$

$p_1(2) = p_0(1)q_{01} + p_1(1)q_{11} + p_2(1)q_{21} + p_3(1)q_{31} = 0{,}148;$

$p_2(2) = p_0(1)q_{02} + p_1(1)q_{12} + p_2(1)q_{22} + p_3(1)q_{32} = 0{,}258;$

$p_3(2) = p_0(1)q_{03} + p_1(1)q_{13} + p_2(1)q_{23} + p_3(1)q_{33} = 0{,}304.$

Аналогично вероятности состояний системы могут быть рассчитаны на моменты времени $t_3, t_4, \ldots$.

Нетрудно убедиться, что сумма вероятностей состояний системы на каждый момент времени равна единице: $p_0(k) + p_1(k) + p_2(k) + p_3(k) = 1$ для $k = 1, 2, \ldots$.

Матрица вероятностей переходов рассматриваемой системы – неразложимая и непериодическая, следовательно, случайный процесс обладает эргодическим свойством, и вероятности состояний системы для стационарного режима (стационарные вероятности) p_0, p_1, p_2, p_3 могут быть найдены из системы линейных алгебраических уравнений (5.10) с учётом нормировочного условия (5.11):

$$\begin{cases} p_0 = p_0 q_{00} + p_1 q_{10} + p_2 q_{20} + p_3 q_{30} = 0{,}5 p_1 + 0{,}5 p_2; \\ p_1 = p_0 q_{01} + p_1 q_{11} + p_2 q_{21} + p_3 q_{31} = 0{,}2 p_0 + 0{,}4 p_3; \\ p_2 = p_0 q_{02} + p_1 q_{12} + p_2 q_{22} + p_3 q_{32} = 0{,}5 p_0 + 0{,}1 p_1 + 0{,}6 p_3; \\ p_3 = p_0 q_{03} + p_1 q_{13} + p_2 q_{23} + p_3 q_{33} = 0{,}3 p_0 + 0{,}4 p_1 + 0{,}5 p_2; \\ p_0 + p_1 + p_2 + p_3 = 1. \end{cases}$$

Решая систему уравнений, получим значения стационарных вероятностей: $p_0 = \dfrac{13}{55} \approx 0{,}236$, $p_1 = \dfrac{9}{55} \approx 0{,}164$, $p_2 = \dfrac{17}{55} \approx 0{,}309$, $p_3 = \dfrac{16}{55} \approx 0{,}291$.

Таким образом, система будет простаивать 23,6% времени, а более 76% времени система будет находиться в рабочем состоянии, причем почти 30% времени (точнее 29,1%) во включённом состоянии будут одновременно находиться оба устройства системы. Среднее число устройств, находящихся одновременно во включённом состоянии, будет равно: $M = p_1 + p_2 + 2p_3 = 1{,}055$, то есть во включённом состоянии находится в среднем одно устройство.

5.3.3. Марковские процессы с непрерывным временем

Для однородного марковского процесса с непрерывным временем вероятности состояний на произвольный момент времени t определяются из системы дифференциальных уравнений:

$$\frac{dp_j(t)}{dt} = \sum_{i=1}^{n} p_i(t)\, g_{ij} \quad (j = \overline{1,n};\ t > 0) \tag{5.12}$$

с учетом начальных условий $p_1(0), \ldots, p_n(0)$.

Для систем обладающих эргодическим свойством, имеет место стационарный режим, для которого, согласно (5.8), вероятности состояний $p_1, \ldots, p_n$ при $t \to \infty$ не зависят от начальных вероятностей и текущего момента времени t, и система дифференциальных уравнений (5.12) для установившегося режима преобразуется в систему линейных алгебраических уравнений:

$$\sum_{i=1}^{n} p_i g_{ij} = 0 \quad (j = \overline{1,n}), \tag{5.13}$$

которая совместно с нормировочным условием (5.11) образует систему, обладающую единственным решением.

Доказательство выражений (5.12) и (5.13).

Рассмотрим однородный марковский процесс с непрерывным временем, который может находиться в одном из n возможных состояний: $\mathbf{E}_1$, …, $\mathbf{E}_n$. Интенсивности переходов q_{ij} заданы в виде матрицы $\mathbf{G} = [g_{ij} \mid i, j = \overline{1,n}]$, в которой диагональные элементы рассчитаны в соответствии с формулой (5.4). Начальные вероятности на момент времени $t = 0$ заданы в виде вектора $P = \{p_1(0),\dots, p_n(0)\}$.

Определим вероятность $p_j(t)$ того, что в момент времени $t > 0$ случайный процесс находится в состоянии $\mathbf{E}_j$.

Придадим времени t малое приращение Δt и найдем вероятность $p_j(t+\Delta t)$ того, что случайный процесс в момент времени $(t+\Delta t)$ окажется в состоянии $\mathbf{E}_j$.

Случайный процесс может оказаться в состоянии $\mathbf{E}_j$ в момент $(t+\Delta t)$ двумя способами:

1) в момент времени t процесс находился в состоянии $\mathbf{E}_j$ и в течение промежутка времени Δt не перешел в другое состояние;

2) в момент времени t процесс находился в состоянии $\mathbf{E}_i$ $(i \ne j)$ и за время Δt совершил переход в состояние $\mathbf{E}_j$.

Вероятность первого способа $p_j^{(1)}(t+\Delta t)$ найдем как произведение вероятности $p_j(t)$ того, что в момент времени t случайный процесс находился в состоянии $\mathbf{E}_j$, на условную вероятность того, что, будучи в $\mathbf{E}_j$, процесс не перешел в другие состояния $\mathbf{E}_i$ $(i = 1, 2,\dots, j-1,\ j+1,\dots,n)$. Эта условная вероятность равна

$$1 - \sum_{\substack{i=1 \\ i \ne j}}^{n} g_{ji} \Delta t\,.$$

Последнее выражение становится очевидным, если вспомнить, что произведение $g_{ji}\Delta t$ с точностью до бесконечно малых высших порядков определяет вероятность перехода случайного процесса из состояния $\mathbf{E}_j$ в состояние $\mathbf{E}_i$ за промежуток времени Δt, а сумма этих вероятностей есть вероятность перехода из состояния $\mathbf{E}_j$ в любое другое состояние, не совпадающее с $\mathbf{E}_j$. Вычитая эту сумму из единицы, получим требуемую вероятность противоположного события.

Таким образом, вероятность первого способа с учетом выражения (5.4) равна

$$p_j^{(1)}(t+\Delta t) = p_j(t)(1+g_{jj}\Delta t) \qquad (j=\overline{1,n}).$$

Аналогично определяется вероятность $p_j^{(2)}(t+\Delta t)$ второго способа оказаться в состоянии $\mathbf{E}_j$ в момент $(t+\Delta t)$: она равна вероятности $p_j(t)$ того, что в момент времени t процесс находился в состоянии $\mathbf{E}_i$, умноженной на вероятность $g_{ij}\Delta t$ перехода за время Δt в состояние $\mathbf{E}_j$: $p_i(t)\,g_{ij}\,\Delta t$. Суммируя эти вероятности по всем возможным состояниям, исключая состояние $\mathbf{E}_j$, получим искомую вероятность:

$$p_j^{(2)}(t+\Delta t) = \sum_{\substack{i=1\\ i\neq j}}^{n} p_i(t)\,g_{ij}\,\Delta t \qquad (j=\overline{1,n}).$$

Применив правило сложения вероятностей, получим вероятность нахождения случайного процесса в состоянии $\mathbf{E}_j$ в момент времени $(t+\Delta t)$:

$$p_j(t+\Delta t) = p_j^{(1)}(t+\Delta t) + p_j^{(2)}(t+\Delta t) = p_j(t)(1+g_{jj}\Delta t) + \sum_{\substack{i=1\\ i\neq j}}^{n} p_i(t)\,g_{ij}\,\Delta t$$

или $p_j(t+\Delta t) - p_j(t) = \sum\limits_{i=1}^{n} p_i(t)\,g_{ij}\,\Delta t \qquad (j=\overline{1,n})$.

Разделим левую и правую части последнего выражения на Δt и перейдем к пределу при $\Delta t \to 0$:

$$\lim_{\Delta t\to 0} \frac{p_j(t+\Delta t) - p_j(t)}{\Delta t} = \sum_{i=1}^{n} p_i(t)\,g_{ij} \qquad (j=\overline{1,n}).$$

Левая часть полученного выражения представляет собой производную по времени от функции $p_j(t)$:

$$\frac{dp_j(t)}{dt} = \sum_{i=1}^{n} p_i(t)\,g_{ij} \qquad (j=\overline{1,n}). \qquad (5.14)$$

Таким образом, получена система дифференциальных уравнений марковского случайного процесса, которая при заданных начальных условиях $P=\{p_1(0),\ldots,\ p_n(0)\}$ позволяет выполнить исследование нестационарного (переходного) режима работы моделируемой системы путем расчёта вероятностей состояний марковского процесса в произвольный момент времени $t>0$.

Для случайных процессов, обладающих эргодическим свойством, имеет место стационарный режим, для которого согласно (5.8) вероятности состояний при $t \to \infty$ не зависят от начальных вероятностей и текущего момента времени *t*. Тогда производные $\frac{dp_j(t)}{dt} = 0$, и система дифференциальных уравнений (5.14) преобразуется в систему линейных алгебраических уравнений (5.13).

Пример. В качестве примера марковского процесса с непрерывным временем рассмотрим модель «гибели и размножения», которая часто встречается в разнообразных практических задачах. Своим названием эта модель обязана биологической задаче об изменении численности популяции и распространении эпидемий, которая формулируется следующим образом.

Рассмотрим развитие некоторой популяции, особи которой могут рождаться и умирать. Положим, что при наличии *i* особей в популяции рождение новых особей происходит с интенсивностью λ_i и с интенсивностью μ_i – особи умирают. Пусть в любой момент времени может происходить рождение или гибель только одной особи, и интервалы времени между двумя моментами рождения и гибели распределены по экспоненциальному закону с параметрами λ_i и μ_i соответственно. Тогда процесс "гибели и размножения" может быть представлен марковским случайным процессом с непрерывным временем (рис.5.1,а), в котором состояние $\mathbf{E}_i$ соответствует наличию *i* особей в популяции (*i*=0, 1, …), причем число состояний может быть конечным или бесконечным. Отметим, что состояние $\mathbf{E}_0$ соответствует вырождению популяции.

Таким образом, марковский процесс называется «**процессом гибели и размножения**», если её граф переходов имеет вид цепочки состояний, в которой каждое состояние (кроме крайних) связано с двумя соседними состояниями, а крайние состояния $\mathbf{E}_0$ и $\mathbf{E}_n$ (в случае конечного числа состояний) или только нулевое состояние $\mathbf{E}_0$ (в случае бесконечного числа состояний) – только с одним соседним состоянием.

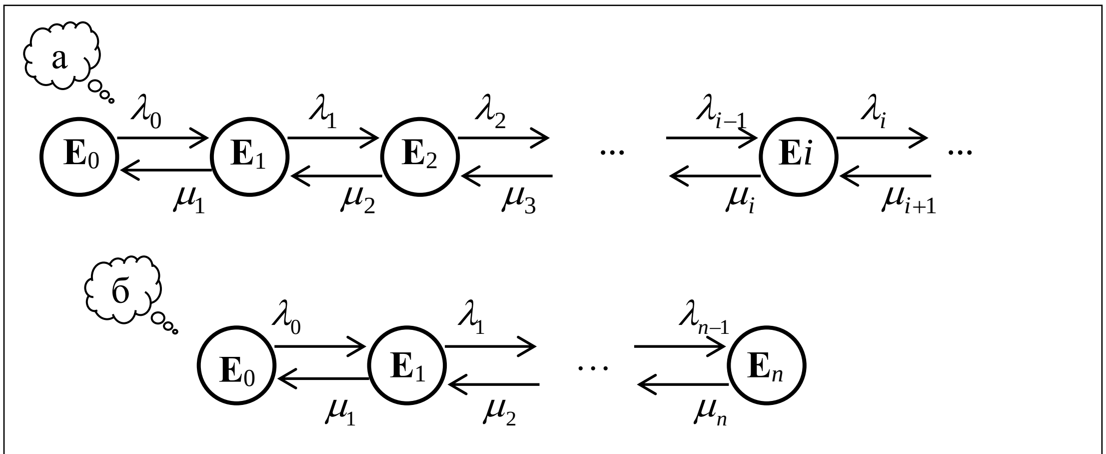


Рис. 5.2. Графы состояний процесса гибели и размножения с бесконечным (а) и конечным (б) числом состояний

Графу переходов процесса гибели и размножения с конечным числом состояний (рис.5.2,б) соответствует матрица интенсивностей переходов:

$$
\mathbf{G} = \begin{array}{c|ccccccc}
\mathbf{E}_i & 0 & 1 & 2 & \dots & n-1 & n \\
\hline
0 & -\lambda_0 & \lambda_0 & 0 & \dots & 0 & 0 \\
1 & \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & \dots & 0 & 0 \\
2 & 0 & \mu_2 & -(\lambda_2+\mu_2) & \dots & 0 & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
n-1 & 0 & 0 & 0 & \dots & -(\lambda_{n-1}+\mu_{n-1}) & \lambda_{n-1} \\
n & 0 & 0 & 0 & \dots & \mu_n & -\mu_n
\end{array}.
$$

Диагональные элементы матрицы определяются из условия (5.4) – сумма элементов каждой строки должна быть равна нулю.

Система линейных алгебраических уравнений для определения стационарных вероятностей может быть составлена по графу переходов или по матрице интенсивностей переходов.

Сформулируем **правила составления уравнений для стационарных вероятностей состояний** марковского процесса с непрерывным временем по графу переходов и по матрице интенсивностей переходов.

Правило 1 (по графу переходов). В левой части каждого уравнения записывается вероятность рассматриваемого состояния, умноженная на сумму интенсивностей переходов из данного состояния во все другие состояния. Правая часть уравнения представляет собой сумму членов, число которых равно числу входящих в данное состояние дуг, и каждый такой член представляет собой произведение интенсивности перехода, соответствующей данной дуге, на вероятность состояния, из которого исходит эта дуга.

Для нашего примера применение правила 1 дает следующую систему линейных алгебраических уравнений:

$$
\begin{cases}
\lambda_0 p_0 = \mu_1 p_1 \\
(\lambda_1+\mu_1) p_1 = \lambda_0 p_0 + \mu_2 p_2 \\
\dots \\
(\lambda_k+\mu_k) p_k = \lambda_{k-1} p_{k-1} + \mu_{k+1} p_{k+1} \\
\dots \\
\mu_n p_n = \lambda_{n-1} p_{n-1} \\
p_0 + p_1 + \dots + p_n = 1
\end{cases},
$$

где последнее уравнение представляет собой нормировочное условие (5.11).

Правило 2 (по матрице интенсивностей переходов). Для каждого столбца матрицы интенсивностей переходов составляется соответствующее уравнение как сумма произведений интенсивностей переходов на стационарную вероятность состояния с номером соответствующей строки, приравненная нулю.

Применение правила 2 для нашего примера дает следующую систему линейных алгебраических уравнений:

$$\begin{cases} -\lambda_0 p_0 + \mu_1 p_1 = 0 \\ -(\lambda_1 + \mu_1) p_1 + \lambda_0 p_0 + \mu_2 p_2 = 0 \\ \ldots \\ -(\lambda_k + \mu_k) p_k + \lambda_{k-1} p_{k-1} + \mu_{k+1} p_{k+1} = 0 \\ \ldots \\ -\mu_n p_n + \lambda_{n-1} p_{n-1} = 0 \\ p_0 + p_1 + \ldots + p_n = 1 \end{cases}.$$

Легко убедиться, что обе системы уравнений эквивалентны.

Решая полученную систему уравнений аналитически или с применением численных методов, можно определить значения $p_0, p_1, \ldots, p_n$ стационарных вероятностей состояний марковского процесса. Кроме того, могут быть рассчитаны другие характеристики исследуемой системы, в частности, среднее число особей в популяции как математическое ожидание случайной величины:

$$M = \sum_{k=1}^{n} k p_k .$$

5.4. Марковские модели систем массового обслуживания

В данном параграфе в качестве примеров применения случайных процессов для изучения свойств систем с дискретным характером функционирования подробно рассматриваются марковские модели систем массового обслуживания (СМО). Примеры представлены в порядке возрастания их сложности, начиная с простейшей одноканальной СМО с однородным потоком заявок без накопителя и заканчивая СМО с накопителями ограниченной ёмкости и приоритетным обслуживанием неоднородного потока заявок.

В каждом примере приводится *описание* исследуемой системы, а также *предположения и допущения*, принятые при построении математической модели и необходимые для того, чтобы протекающий в системе случайный процесс был марковским. Разработка марковской модели исследуемой системы в терминах случайных процессов предполагает выполнение следующих этапов:

- *кодирование состояний* случайного процесса;
- построение *размеченного графа переходов*;

- формирование *матрицы интенсивностей переходов*;
- составление *системы линейных алгебраических уравнений* для расчёта стационарных вероятностей состояний марковского процесса.

Матрица интенсивностей переходов может использоваться для задания системы линейных алгебраических уравнений в матричном виде при компьютерном расчёте стационарных вероятностей.

При исследовании различного рода *реальных* систем, моделями которых служат СМО, вряд ли кого-то интересуют вероятности состояний. Гораздо больший интерес представляют такие характеристики СМО, как длина очереди заявок перед обслуживающим прибором, время ожидания и время пребывания заявок в системе, загрузка и коэффициент простоя системы, доля потерянных заявок и т.д., значения которых могут быть рассчитаны по найденным значениям стационарных вероятностей состояний. Поэтому ниже особое внимание уделяется математическим зависимостям, позволяющим рассчитать в каждом конкретном примере наиболее важные характеристики функционирования исследуемых систем. Максимально подробно процесс получения таких зависимостей изложен в нескольких первых рассматриваемых ниже примерах. На основе этих зависимостей в некоторых примерах проводится анализ свойств исследуемой системы. Для остальных примеров подобный анализ рекомендуется читателю выполнить самостоятельно.

В первом примере (п.5.4.1) одноканальной СМО без накопителя представлена диаграмма функционирования исследуемой системы, с помощью которой показано, что протекающий в системе случайный процесс при сформулированных предположениях и допущениях (а это, прежде всего, экспоненциальный характер процессов поступления и обслуживания заявок) является марковским. Аналогично, и для остальных систем можно показать, что протекающие в них случайные процессы при сформулированных предположениях и допущениях являются марковскими.

В некоторых случаях на основе марковских моделей могут быть получены математические выражения для расчёта стационарных вероятностей состояний в явном виде без применения методов численного анализа. В частности, такие результаты представлены ниже для одноканальной и многоканальной СМО без накопителя (с отказами) и одноканальной СМО с накопителем неограниченной и ограниченной ёмкости.

5.4.1. Одноканальная СМО без накопителя (M/M/1/0)

Рассмотрим простейшую одноканальную систему массового обслуживания (СМО) с отказами, в которую поступает случайный поток заявок, задерживаемых в приборе на случайное время (рис.5.3,а). Поскольку перед обслуживающим прибором нет накопителя, то заявка, поступившая в систему и заставшая прибор занятым, получает отказ в обслуживании и теряется. Таким образом, в системе, кроме входящего потока заявок с интенсивностью λ, образуются еще два потока: выходящий поток обслуженных в приборе заявок с интенсивностью λ' и поток необслуженных заявок (получивших отказ в обслуживании) с интенсивностью λ''. Очевидно, что $\lambda' + \lambda'' = \lambda$.

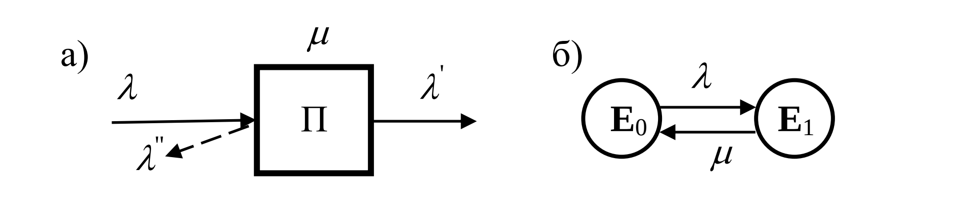


Рис. 5.3. СМО с отказами (а) и её граф переходов (б)

1. Описание системы.

1.1. Система содержит один обслуживающий прибор (П), то есть является *одноканальной*.

1.2. В систему поступает один класс заявок, то есть поток заявок *однородный.*

1.3. В приборе происходит задержка (обслуживание) поступающих в систему заявок на некоторое случайное время.

1.4. Перед прибором не предусмотрены места для ожидания заявок, то есть в системе отсутствует накопитель.

2. Предположения и допущения.

2.1. Поступающие в систему заявки образуют *простейший* поток с интенсивностью λ.

2.2. Длительность обслуживания заявок в приборе распределена по *экспоненциальному* закону с интенсивностью $\mu = 1/b$, где b – средняя длительность обслуживания.

2.3. Дисциплина буферизации – *с отказами*: заявка, поступившая в систему и заставшая прибор занятым обслуживанием другой заявки, теряется.

2.4. Дисциплина обслуживания – *в естественном порядке*: заявка, поступившая в систему и заставшая прибор свободным, принимается на обслуживание.

Очевидно, что в СМО с отказами всегда будет существовать установившийся режим, поскольку даже при больших значениях нагрузки ($y >> 1$) число заявок в системе не может вырасти до бесконечности. Это обусловлено тем, что с ростом нагрузки увеличивается доля заявок, получающих отказ в обслуживании.

3. Кодирование состояний случайного процесса.

В качестве параметра, описывающего состояние случайного процесса, будем рассматривать количество заявок k, находящихся в СМО. Очевидно, что система в любой момент времени может находиться в одном из двух состояний:

$\mathbf{E}_0$: $k = 0$ – в системе нет заявок (прибор простаивает);

$\mathbf{E}_1$: $k = 1$ – в системе (на обслуживании в приборе) находится 1 заявка (прибор работает).

4. Размеченный граф переходов случайного процесса (рис.5.3,б).

В процессе функционирования рассматриваемой системы в один и тот же момент времени может наступить только одно из двух возможных событий, которые приводят к изменению состояния случайного процесса, протекающего в системе.

1. *Поступление заявки в систему* с интенсивностью λ. При этом:

•если случайный процесс находится в состоянии $\mathbf{E}_0$ (прибор простаивает), то произойдет переход в состояние $\mathbf{E}_1$ (начнется обслуживание поступившей заявки), причем *интенсивность перехода совпадает с интенсивностью поступления* заявок в систему λ;

•если же случайный процесс находится в состоянии $\mathbf{E}_1$ (прибор работает), то состояние $\mathbf{E}_1$ случайного процесса не изменится, что будет соответствовать отказу в обслуживании поступившей заявке.

Таким образом, переход из состояния $\mathbf{E}_0$ в состояние $\mathbf{E}_1$ происходит с интенсивностью λ.

2. *Завершение обслуживания заявки*, находящейся в приборе. Это событие может наступить только в том случае, если в приборе на обслуживании находится заявка, то есть случайный процесс находится в состоянии $\mathbf{E}_1$. При этом происходит переход в состояние $\mathbf{E}_0$, причем *интенсивность перехода совпадает с интенсивностью обслуживания* заявки в приборе μ.

5. Диаграммы функционирования системы.

Рассмотрим диаграммы функционирования системы и покажем, что случайный процесс, протекающий в системе, при сформулированных выше предположениях является марковским.

На рис.5.4 показаны диаграммы следующих процессов:

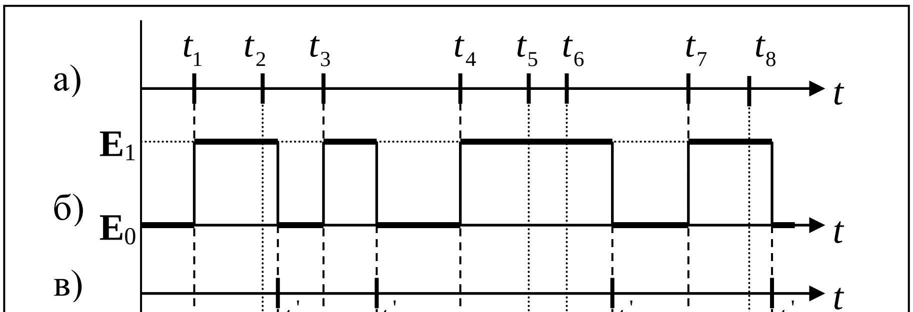
а) **поступления** в СМО заявок, интервалы между которыми в случае простейшего потока распределены по экспоненциальному закону;

б) **перехода** из состояния $\mathbf{E}_0$ в состояние $\mathbf{E}_1$ и обратно, в которых может находиться система, при этом время нахождения случайного процесса в состоянии $\mathbf{E}_1$ равно длительности обслуживания заявки в приборе, которая представляет собой случайную величину, распределенную по экспоненциальному закону;

в) **выхода** из системы обслуженных заявок в моменты времени t_1', t_3', t_4', t_7';

г) **выхода** из системы необслуженных заявок, получивших отказ из-за занятости прибора в моменты времени t_2', t_5', t_6', t_8';

д) **формирования** интервалов времени между соседними переходами случайного процесса.



Как показано в п. 5.1.2, дискретный случайный процесс с непрерывным временем будет марковским, если интервалы между соседними переходами распределены по *экспоненциальному* закону.

В нашем случае, интервал τ_1 представляет собой интервал между поступающими заявками, который для простейшего потока имеет экспоненциальное распределение. Интервалы $\tau_2, \tau_4, \tau_6, \tau_8$, как видно из диаграммы, представляют собой время нахождения случайного процесса в состоянии **E**$_1$, равное длительности обслуживания заявки в приборе, которая распределена по экспоненциальному закону. Таким образом, интервалы $\tau_1, \tau_2, \tau_4, \tau_6, \tau_8$ *распределены по экспоненциальному закону* и, следовательно, удовлетворяют сформулированному выше условию.

Рассмотрим теперь выделенные интервалы τ_3, τ_5, τ_7. Каждый из этих интервалов представляет собой промежуток времени от момента завершения обслуживания некоторой заявки до момента поступления новой заявки, принимаемой на обслуживание в приборе. В п.5.1.2 сформулировано замечательное *свойство экспоненциального распределения*, которое гласит, что в случае экспоненциального распределения интервалов времени между двумя событиями *интервал времени от любого случайного момента до момента наступления очередного события* имеет *такое же экспоненциальное распределение с тем же параметром*. В соответствии с этим свойством интервалы τ_3, τ_5, τ_7 имеют *экспоненциальное распределение* с параметром λ и, следовательно, также удовлетворяют сформулированному выше условию для марковского процесса.

Таким образом, случайный процесс, протекающий в системе с простейшим потоком заявок и экспоненциальным обслуживанием, является марковским.

6. Матрица интенсивностей переходов.

Графу переходов (рис.5.3) соответствует **матрица интенсивностей переходов**:

$$\mathbf{G} = \begin{array}{c|cc} \mathbf{E}_i & 0 & 1 \\ \hline 0 & -\lambda & \lambda \\ 1 & \mu & -\mu \end{array} .$$

Действительно, переход из состояния $\mathbf{E}_0$ в состояние $\mathbf{E}_1$ соответствует поступлению заявки в систему с интенсивностью λ, а переход из состояния $\mathbf{E}_1$ в состояние $\mathbf{E}_0$ соответствует завершению обслуживания заявки в приборе с интенсивностью μ.

Диагональные элементы матрицы определяются из условия (5.4) – сумма элементов каждой строки должна быть равна нулю.

7. Система уравнений.

Система уравнений для определения стационарных вероятностей, составленная по графу переходов с применением *правила 1*, имеет вид:

$$\begin{cases} \lambda p_0 = \mu p_1 \\ \mu p_1 = \lambda p_0 \\ p_0 + p_1 = 1 \end{cases},$$

где последнее уравнение представляет собой нормировочное условие (5.10).

Учитывая, что первое и второе уравнение одинаковы (или, как говорят математики, линейно зависимы) и удаляя одно из них, окончательно получим:

$$\begin{cases} \lambda p_0 = \mu p_1 \\ p_0 + p_1 = 1 \end{cases}.$$

Решая эту систему, получим следующие значения стационарных вероятностей состояний марковского процесса:

$$p_0 = \frac{\mu}{\lambda + \mu} = \frac{1}{1+y}; \qquad p_1 = \frac{\lambda}{\lambda + \mu} = \frac{y}{1+y},$$

где $y = \lambda / \mu$ - нагрузка системы.

8. Расчет характеристик СМО.

Для расчета характеристик СМО можно воспользоваться следующими математическими зависимостями, вытекающими из зависимостей (3.6) – (3.18):

1) нагрузка: $y = \lambda / \mu = \lambda b$ (по определению);

2) загрузка определяется как *вероятность работы прибора*: $\rho = p_1$ и не совпадает с нагрузкой даже в случае $y < 1$, что характерно для систем с отказами и потерями заявок, причём всегда $\rho < y$;

3) коэффициент простоя системы определяется как вероятность отсутствия заявок в системе или, по определению, через загрузку системы: $\eta = p_0 = 1 - \rho$;

4) среднее число заявок в системе: $m = p_1 = \rho$, определяемое как математическое ожидание случайной величины: в системе может находиться либо ноль заявок с вероятностью p_0, либо одна заявка с вероятностью p_1, тогда среднее число заявок равно $m = 0 \cdot p_0 + 1 \cdot p_1 = p_1$;

5) вероятность потери заявок в результате отказа в обслуживании из-

за занятости прибора в соответствии с (3.18) совпадает с вероятностью того, что система занята обслуживанием заявок:

$$\pi = \pi_n = 1 - \frac{\rho}{y} K = 1 - \frac{p_1}{y} = \frac{y}{1+y} = p_1,$$

где учтено, что для рассматриваемой СМО без накопителя $p_1 = \dfrac{y}{1+y}$;

6) производительность системы: $\lambda' = (1-\pi)\lambda$, определяемая как интенсивность потока обслуженных заявок на выходе системы;

7) интенсивность потока не обслуженных заявок, то есть получивших отказ: $\lambda'' = \pi\lambda$;

8) среднее время пребывания заявок в системе: $u = m/\lambda' = b$ (определяется по формуле Литтла (3.15) и, как следовало ожидать, равно средней длительности обслуживания заявок; отметим, что в формуле Литтла используется интенсивность λ' потока *обслуженны*х заявок, а не входящего потока λ).

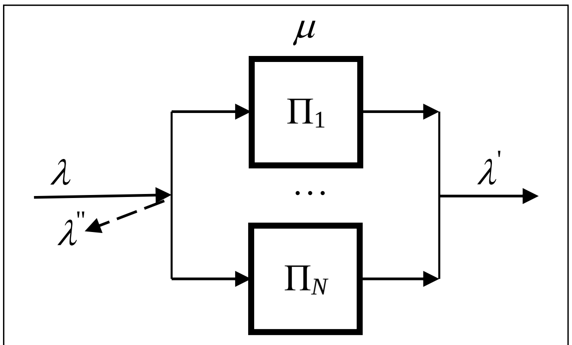


Рис. 5.6. Многоканальная СМО с отказами

9. Анализ свойств системы.

Анализ полученных зависимостей (рис.5.5) показывает, что с ростом нагрузки коэффициент простоя системы, равный p_0, уменьшается, а загрузка системы, определяемая как вероятность p_1 того, что прибор ра-ботает, (а также среднее число зая-вок в системе и вероятность отказа) увеличивается, причем их сумма всегда равна единице. При $y \to \infty$ коэф-фициент простоя $\eta \to 0$, в то время как загрузка $\rho \to 1$. Заметим также, что нагрузка системы определяется через стационарные вероятности как отношение вероятности работы системы к вероятности простоя: $y = p_1 / p_0$ (что легко может быть получено из выражений для p_0 и p_1) или, что то же самое, через загрузку и коэффициент простоя: $y = \rho/\eta$.

5.4.2. Многоканальная СМО без накопителя (M/M/N/0)

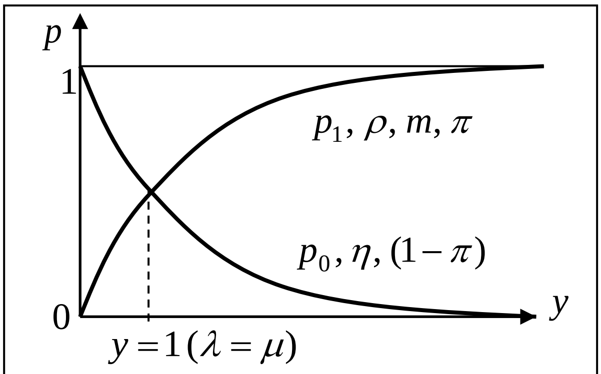


Рис.5.5. Характеристики СМО

Рассмотрим многоканальную систему массового обслуживания (СМО) с отказами, в которую поступает случайный поток заявок, задержи-ваемых в приборе на случайное время (рис.5.6). Заявка, поступившая в систему и заставшая прибор занятым, получает отказ в обслуживании и теряется. Таким образом, в системе, кроме входящего потока заявок с интенсивностью λ, образуются еще два потока: выходящий поток обслу-женных в приборе заявок с интенсив-ностью $\lambda^{'}$ и поток необслуженных заявок (получивших отказ в обслу-живании) с интенсивностью $\lambda^{''}$. Очевидно, что $\lambda^{'} + \lambda^{''} = \lambda$.

1. Описание системы.

1.1. Система содержит *N* обслуживающих приборов $П_1,\ldots,П_N$, то есть является *многоканальной*.

1.2. В систему поступает один класс заявок (поток *однородный*).

1.3. Все приборы идентичны, то есть любая заявка может быть обслужена любым прибором за одно и то же случайное время.

1.4. В системе отсутствует накопитель.

2. Предположения и допущения.

2.1. Поступающие в систему заявки образуют *простейший* поток с интенсивностью λ.

2.2. Длительность обслуживания заявок в любом приборе распределена по экспоненциальному закону с интенсивностью $\mu = 1/b$, где b – средняя длительность обслуживания заявок в приборе.

2.3. Дисциплина буферизации – *с отказами*: заявка, поступившая в систему и заставшая все приборы занятыми обслуживанием других заявок, теряется.

2.4. Дисциплина обслуживания – *в естественном порядке*: заявка, поступившая в систему принимается на обслуживание, если есть хотя бы один свободный прибор. Если заявка застала свободными несколько приборов, то она направляется в один из них случайным образом.

3. Кодирование состояний случайного процесса.

В качестве параметра, описывающего состояние случайного процесс-са, как и ранее, будем рассматривать количество заявок *k*, находящихся в СМО. При этом система в любой момент времени может находиться в одном из $(N+1)$ состояний:

E$_0$: $k = 0$ – в системе нет заявок (система простаивает);

E$_1$: $k = 1$ – в системе находится 1 заявка (один прибор работает, остальные – простаивают);

E$_2$: $k = 2$ – в системе находится 2 заявки (два прибора работают, остальные – простаивают);

…

E$_N$: $k = N$ – в системе находится N заявок (все приборы работают).

4. Размеченный граф переходов случайного процесса представлен на рис.5.7.

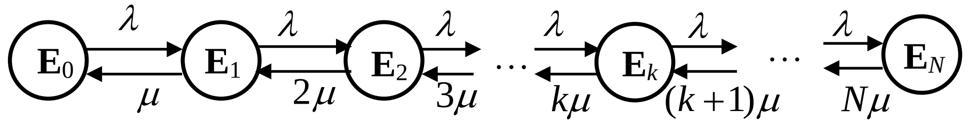


Рис. 5.7. Граф переходов марковского процесса

В один и тот же момент времени в системе может произойти только одно из двух событий, которые приводят к изменению состояния случайного процесса.

1. *Поступление заявки в систему* с интенсивностью λ. При этом:

•если случайный процесс находится в состоянии $\mathbf{E}_k$, причем $k < N$, то произойдет переход в состояние $\mathbf{E}_{k+1}$ (начнется обслуживание поступившей заявки в одном из свободных приборов), причем интенсивность перехода равна интенсивности поступления λ;

•если же случайный процесс находится в состоянии $\mathbf{E}_N$ (все приборы заняты обслуживанием заявок), то состояние $\mathbf{E}_N$ случайного процесса не изменится, что будет соответствовать отказу в обслуживании поступившей заявке.

Таким образом, переход из состояний $\mathbf{E}_k$ в состояние $\mathbf{E}_{k+1}$ (при $k < N$) происходит с интенсивностью λ.

2. *Завершение обслуживания заявки* в одном из приборов с интенсивностью μ.

Это событие может наступить только в том случае, если в системе на обслуживании находится хотя бы одна заявка, то есть случайный процесс находится в состояниях $\mathbf{E}_1$, $\mathbf{E}_2$, …, $\mathbf{E}_N$,. При этом случайный процесс переходит соответственно в состояния $\mathbf{E}_0$, $\mathbf{E}_1$, …, $\mathbf{E}_{N-1}$, причём интенсивности перехода различны. Действительно, если в системе обслуживается только одна заявка (состояние $\mathbf{E}_1$), то интенсивность перехода в состояние $\mathbf{E}_0$ равна μ. Если же в системе обслуживается две заявки (состояние $\mathbf{E}_2$), то есть работают два прибора, то переход случайного процесса в состояние $\mathbf{E}_1$ возможен либо в результате завершения обслуживания заявки в первом приборе с интенсивностью μ, либо в результате завершения обслуживания заявки во втором приборе с такой же интенсивностью μ, причём вероятность завершения обслуживания заявок в обоих приборах в один и тот же момент времени равна нулю. Таким образом, интенсивность перехода из состояния $\mathbf{E}_2$ в состояние $\mathbf{E}_1$ будет равна 2μ (как сумма интенсивностей двух рассмотренных способов).

В общем случае, если в многоканальной системе на обслуживании находится $k = 1,2,...,N$ заявок (случайный процесс находится в состоянии $\mathbf{E}_k$), то интенсивность перехода в состояние $\mathbf{E}_{k-1}$ будет равна $k\mu$.

По аналогии с предыдущим примером (п.5.4.1) здесь и в последующих примерах можно показать, что случайный процесс, протекающий в системе, при сформулированных предположениях является марковским.

5. Матрица интенсивностей переходов.

Графу переходов (рис.5.7) соответствует матрица интенсивностей переходов:

$$\mathbf{G} = \begin{array}{c|cccccc}
\mathbf{E}_i & 0 & 1 & 2 & \dots & N-1 & N \\
\hline
0 & -\lambda & \lambda & 0 & \dots & 0 & 0 \\
1 & \mu & -(\lambda+\mu) & \lambda & \dots & 0 & 0 \\
2 & 0 & 2\mu & -(\lambda+2\mu) & \dots & 0 & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
N-1 & 0 & 0 & 0 & \dots & -(\lambda+(N-1)\mu) & \lambda \\
N & 0 & 0 & 0 & \dots & N\mu & -N\mu
\end{array}.$$

Диагональные элементы матрицы определяются из условия (5.4) – сумма элементов каждой строки должна быть равна нулю.

6. Система уравнений.

Система уравнений для определения стационарных вероятностей имеет вид:

$$\begin{cases}
\lambda p_0 = \mu p_1 \\
(\lambda+\mu) p_1 = \lambda p_0 + 2\mu p_2 \\
\dots \\
(\lambda + k\mu) p_k = \lambda p_{k-1} + (k+1)\mu p_{k+1} \\
\dots \\
N\mu p_N = \lambda p_{N-1} \\
p_0 + p_1 + \dots + p_N = 1
\end{cases}.$$

Используя метод математической индукции можно показать, что:

$$p_k = \frac{y^k}{k!} p_0 \qquad (k = \overline{0, N}),$$

где $y = \lambda b$ – нагрузка системы.

Подставляя полученное выражение в последнее уравнение системы линейных алгебраических уравнений, найдем вероятность простоя системы:

$$p_0 = \frac{1}{\sum_{i=0}^{N} \frac{y^i}{i!}}.$$

Тогда стационарные вероятности состояний марковского случайного процесса, протекающего в многоканальной СМО с отказами:

$$\boxed{p_k = \frac{\frac{y^k}{k!}}{\sum_{i=0}^{N} \frac{y^i}{i!}} \qquad (k = \overline{0, N})}.$$

Из последнего выражения при $N = 1$, как частный случай, вытекает результат, полученный в предыдущем примере для одноканальной СМО с отказами.

Задание на самостоятельную работу: *проверить полученные выражения, используя метод математической индукции.*

7. Расчет характеристик СМО.

Для расчета характеристик СМО можно воспользоваться следующими математическими зависимостями:

1) нагрузка: $y = \lambda / \mu = \lambda b$ (по определению);

2) загрузка: $\rho = \dfrac{1}{N}\sum_{k=0}^{N} k p_k$, учитывающая долю $\left(\dfrac{k}{N}\right)$ работающих приборов; действительно, система загружена полностью, когда работают все приборы, если же из 10 приборов работает один, то система загружена на 10%, если работают 5 приборов, то система загружена на 50%;

3) коэффициент простоя системы: $\eta = \dfrac{1}{N}\sum_{k=0}^{N}(N-k)p_k = 1-\rho$;

4) среднее число заявок в системе, равное среднему числу работающих приборов: $m = \sum_{k=1}^{N} k p_k = N\rho$;

5) среднее число простаивающих приборов: $\hat{N} = N - m$;

6) вероятность отказа в обслуживании, определяемая как вероятность того, что все приборы заняты обслуживанием заявок: $\pi = p_N = \dfrac{y^N}{N!} \Big/ \sum_{i=0}^{N} \dfrac{y^i}{i!}$;

Задание на самостоятельную работу: *доказать последнее выражение для вероятности потери заявок, подставив полученное выше выражение для стационарных вероятностей состояний в формулу (3.18).*

7) производительность системы, определяемая как интенсивность потока обслуженных заявок: $\lambda' = \lambda(1-\pi)$;

8) интенсивность потока не обслуженных заявок, то есть получивших отказ: $\lambda'' = \lambda\pi$;

9) среднее время пребывания заявок в системе: $u = m / \lambda' = b$.

8. Анализ свойств системы.

Анализ свойств многоканальной СМО без накопителя показывает, что с увеличением нагрузки уменьшается вероятность простоя системы и увеличивается загрузка системы, а вместе с ней число работающих приборов и вероятность отказа.

5.4.3. Одноканальная СМО с накопителем ограниченной емкости (M/M/1/r)

1. Описание системы.

1.1. Система (рис.5.8) содержит один обслуживающий прибор (П), то есть является *одноканальной*.

1.2. Поток поступающих в систему заявок *однородный.*

1.3. Длительность обслуживания заявок в приборе – величина *случайная*.

1.4. Перед прибором имеется r мест для заявок, ожидающих обслуживания и образующих очередь, то есть в системе имеется накопитель *ограниченной* ёмкости: $r < \infty$.

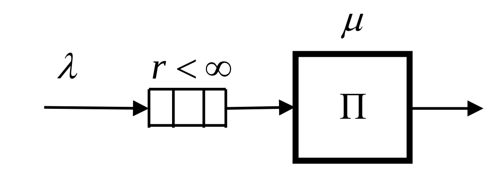


Рис. 5.8. СМО с накопителем ограниченной ёмкости

2. Предположения и допущения.

2.1. Поступающие в систему заявки образуют *простейший* поток с интенсивностью λ.

2.2. Длительность обслуживания заявок в приборе распределена по *экспоненциальному* закону с интенсивностью $\mu = 1/b$, где b – средняя длительность обслуживания заявок в приборе.

2.3. Дисциплина буферизации – *с потерями*: заявка, поступившая в систему и заставшая накопитель заполненным, теряется.

2.4. Дисциплина обслуживания – *в порядке поступления* по правилу «первым пришел – первым обслужен» (FIFO).

В СМО с накопителем ограниченной ёмкости всегда существует установившийся режим, поскольку длина очереди не будет расти до бесконечности даже при больших значениях нагрузки.

3. Кодирование состояний марковского процесса.

В качестве параметра, описывающего состояние марковского процесса, будем рассматривать количество заявок *k*, находящихся в СМО (в приборе и в накопителе). Тогда марковский процесс в любой момент времени может находиться в одном из следующих $(r+2)$-х состояний:

$\mathbf{E}_0$: $k = 0$ – в системе нет ни одной заявки;

$\mathbf{E}_1$: $k = 2$ – в системе находится 1 заявка на обслуживании в приборе;

$\mathbf{E}_2$: $k = 2$ – в системе находятся 2 заявки: одна – на обслуживании в приборе и вторая ожидает в накопителе;

…

$\mathbf{E}_{r+1}$: $k = r+1$ – в системе находятся $(r+1)$ заявок: одна – на обслуживании в приборе и r – в накопителе.

4. Размеченный граф переходов случайного процесса представлен на рис.5.9.

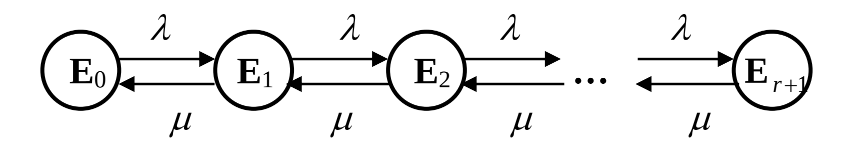


Рис. 5.9. Граф переходов марковского процесса

В один и тот же момент времени в системе может произойти только одно событие:

- поступление заявки с интенсивностью λ, что соответствует увеличению на единицу числа заявок в системе и переходу случайного процесса в состояние с номером на единицу больше;

•завершение обслуживания заявки в приборе с интенсивностью μ, что соответствует уменьшению числа заявок в системе и переходу случайного процесса в состояние с номером на единицу меньше.

Задание на самостоятельную работу: *по графу переходов рис.5.9 построить матрицу интенсивностей переходов.*

5. Система уравнений.

Составим по графу переходов систему уравнений для определения стационарных вероятностей:

$$\begin{cases} \lambda p_0 = \mu p_1 \\ (\lambda+\mu) p_1 = \lambda p_0 + \mu p_2 \\ (\lambda+\mu) p_2 = \lambda p_1 + \mu p_3 \\ \dots \\ \mu p_{r+1} = \lambda p_r \\ \sum\limits_{k=0}^{r+1} p_k = 1 \end{cases}.$$

Используя метод математической индукции можно показать, что

$$p_k = y^k p_0 \quad (k = \overline{0, r+1}),$$

где $y = \lambda b$ – нагрузка системы.

Подставляя полученное выражение в последнее уравнение системы линейных алгебраических уравнений, найдем вероятность простоя системы в зависимости от нагрузки:

$$p_0 = \frac{1}{\sum\limits_{k=0}^{r+1} y^k} = \begin{cases} \dfrac{1-y}{1-y^{r+2}}, & y \neq 1 \\ \dfrac{1}{r+2}, & y = 1 \end{cases}.$$

Тогда стационарные вероятности состояний $p_k \quad (k = \overline{0, r+1})$:

$$p_k = \begin{cases} \dfrac{y^k(1-y)}{1-y^{r+2}}, & y \neq 1 \\ \dfrac{y^k}{r+2}, & y = 1 \end{cases}.$$

Задание на самостоятельную работу: *вывести представленные математические зависимости.*

5. Расчет характеристик СМО.

Характеристики СМО при найденных значениях стационарных вероятностей состояний случайного процесса могут быть рассчитаны по следующим формулам:

1) нагрузка $y = \lambda / \mu = \lambda b$;

2) загрузка $\rho = \sum_{k=1}^{r+1} p_k = 1 - p_0$;

3) коэффициент простоя системы $\eta = p_0 = 1 - \rho$;

4) среднее число заявок в очереди $l = \sum_{k=2}^{r+1} (k-1) p_k$;

5) среднее число заявок в системе $m = \sum_{k=1}^{r+1} k p_k = l + \rho$;

6) вероятность потери заявок $\pi = p_{r+1}$;

Задание на самостоятельную работу: *используя выражение (3.18), доказать, что вероятность потери заявок равна вероятности того, что система заполнена, то есть в накопителе нет свободных мест для вновь поступающих заявок.*

7) производительность системы (интенсивность потока обслуженных заявок) $\lambda^{'} = \lambda(1-\pi)$;

8) интенсивность потока потерянных заявок $\lambda^{''} = \lambda \pi$;

9) среднее время ожидания заявок $w = l / \lambda^{'}$;

10) среднее время пребывания заявок $u = m / \lambda^{'} = w + b$.

5.4.4. Одноканальная СМО с накопителем неограниченной емкости (M/M/1)

1. Описание системы (рис.5.10).

1.1. Система – *одноканальная* – с одним обслуживающим прибором.

1.2. Поток заявок *однородный.*

1.3. В приборе происходит задержка поступающих в систему заявок на некоторое *случайное* время.

1.4. В системе имеется накопитель *неограниченной* ёмкости: $r = \infty$, то есть любая заявка, поступившая в систему, найдет место для ожидания в очереди и не будет потеряна.

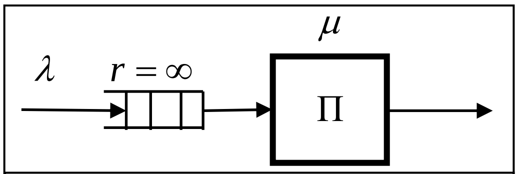


Рис. 5.10. СМО с накопителем неограниченной ёмкости

2. Предположения и допущения.

2.1. Поступающие в систему заявки образуют *простейший* поток с интенсивностью λ.

2.2. Длительность обслуживания заявок в приборе распределена по *экспоненциальному* закону с интенсивностью $\mu = 1/b$, где b – средняя длительность обслуживания заявок в приборе.

2.3. Дисциплина буферизации отсутствует, поскольку накопитель имеет неограниченную ёмкость.

2.4. Дисциплина обслуживания – *в порядке поступления* по правилу «первым пришел – первым обслужен» (FIFO).

2.5. Нагрузка системы совпадает с загрузкой, причём выполняется условие: $y = \rho < 1$, то есть система работает в установившемся режиме без перегрузок. При $y > 1$, в отличие от предыдущих моделей, в СМО устанавливается режим перегрузок.

3. Кодирование состояний марковского процесса.

В качестве параметра, описывающего состояние марковского процесса, как и в предыдущем примере, будем рассматривать количество заявок k, находящихся в СМО (в приборе и в накопителе). Поскольку в системе в произвольный момент времени может находиться любое сколь угодно большое число заявок, то количество состояний марковского процесса равно бесконечности:

$\mathbf{E}_0$: k = 0 – в системе нет ни одной заявки;

$\mathbf{E}_1$: k = 1 – в системе находится 1 заявка (на обслуживании в приборе);

$\mathbf{E}_2$: k = 2 – в системе находятся 2 заявки (одна – на обслуживании в приборе и вторая ожидает в накопителе);

…

$\mathbf{E}_k$: k – в системе находятся k заявок (одна – на обслуживании в приборе и $(k-1)$ – в накопителе).

…

4. Размеченный граф переходов случайного процесса представлен на рис.5.11.

В один и тот же момент времени может происходить только одно событие: поступление заявки в систему с интенсивностью λ или завершение обслуживания заявки с интенсивностью μ. Размеченный граф переходов содержит бесконечное число состояний.

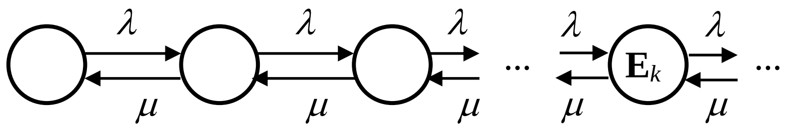


Рис. 5.11. Граф переходов марковского процесса

5. Система уравнений.

Не выписывая матрицу интенсивностей переходов, составим по графу переходов систему уравнений для определения стационарных вероятностей:

$$\begin{cases} \lambda p_0 = \mu p_1 \\ (\lambda+\mu) p_1 = \lambda p_0 + \mu p_2 \\ (\lambda+\mu) p_2 = \lambda p_1 + \mu p_3 \\ \ldots \\ (\lambda+\mu) p_k = \lambda p_{k-1} + \mu p_{k+1} \\ \ldots \\ p_0 + p_1 + p_2 + \ldots + p_k + \ldots = 1 \end{cases}$$

Несмотря на то, что система содержит *бесконечное число уравнений* и, соответственно, *бесконечное число неизвестных*, нетрудно методом математической индукции получить аналитическое решение в явном виде для расчета вероятностей состояний одноканальной экспоненциальной СМО с однородным потоком заявок и накопителем неограниченной ёмкости при условии, что нагрузка системы $y < 1$:

$$p_k = y^k (1-y) = \rho^k (1-\rho) \quad (k = 0, 1, 2, \ldots),$$

где $\rho = \lambda / \mu$ – загрузка системы, совпадающая с нагрузкой, причём $\rho < 1$, что гарантирует существование установившегося режима в системе.

Таким образом, вероятность нахождения марковского процесса в состоянии $\mathbf{E}_k$ или, что то же самое, вероятность того, что в произвольный момент времени в системе находится k заявок, распределена по геометрическому закону с параметром, равным загрузке (нагрузке) системы.

6. Расчет характеристик СМО.

Для расчета характеристик СМО можно воспользоваться следующими математическими зависимостями:

1) нагрузка $y = \lambda / \mu = \lambda b$;

2) загрузка $\rho = 1 - p_0 = \lambda b$ и совпадает с нагрузкой;

3) коэффициент простоя системы $\eta = p_0 = 1 - \rho$;

4) среднее число заявок в очереди $l = \sum_{k=1}^{\infty} (k-1) p_k = \dfrac{\rho^2}{1-\rho}$;

5) среднее число заявок в системе $m = \sum_{k=0}^{\infty} k\, p_k = \dfrac{\rho}{1-\rho}$;

6) вероятность потери заявок $\pi = 0$;

7) производительность системы при отсутствии потерь совпадает с интенсивностью поступления заявок в систему: $\lambda' = \lambda$;

8) интенсивность потерянных заявок $\lambda'' = 0$;

9) среднее время ожидания заявок $w = \dfrac{l}{\lambda} = \dfrac{\rho b}{1-\rho}$;

10) среднее время пребывания заявок $u = w + b$ или $u = \dfrac{m}{\lambda} = \dfrac{b}{1-\rho}$.

Полученные выражения совпадают с формулами для расчёта характеристик экспоненциальной СМО, представленными в п.4.1.1.

Детальный анализ свойств таких систем выполнен в п.4.1.5.

5.4.5. Многоканальная СМО накопителем ограниченной ёмкости (M/M/2/1)

Рассмотрим многоканальную СМО с отказами, в которую поступает однородный поток заявок. Все приборы системы являются идентичными, и поступившая в систему заявка занимает любой свободный прибор. Заявка, поступившая в систему и заставшая все приборы занятыми, заносится в накопитель ограниченной ёмкости, если он не заполнен до предела. В противном случае, заявка получает отказ и покидает систему не обслуженной. Таким образом, в системе, кроме входящего потока заявок с интенсивностью λ, образуются еще два потока заявок: поток обслуженных заявок с интенсивностью λ' и поток необслуженных заявок (получивших отказ в обслуживании) с интенсивностью λ''. Очевидно, что $\lambda' + \lambda'' = \lambda$.

1. Описание системы (рис.5.12).

1.1. Система *двухканальная* - содержит два обслуживающих прибора.

1.2. В систему поступает *однородный* поток заявок.

1.3. Приборы – *идентичные*, то есть время обслуживания заявок в приборах одинаково.

1.4. Перед прибором имеется накопитель *единичной ёмкости*: $r = 1$.

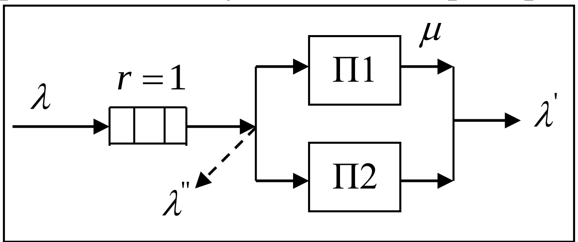


Рис. 5.12. Двухканальная СМО

2. Предположения и допущения.

2.1. Поступающие в систему заявки образуют *простейший* поток с интенсивностью λ.

2.2. Длительность обслуживания заявок в приборе распределена по *экспоненциальному* закону с интенсивностью $\mu = 1/b$, где b – средняя длительность обслуживания заявок в приборе.

2.3. Дисциплина буферизации – *с потерями*: заявка, поступившая в систему и заставшая накопитель заполненным, теряется.

2.4. Дисциплина обслуживания – *в естественном порядке*: заявка, поступившая в систему и заставшая прибор свободным, принимается на обслуживание.

3. Кодирование состояний случайного процесса.

В качестве параметра, описывающего состояние случайного процесса, будем рассматривать количество заявок k, находящихся в СМО. Система в любой момент времени может находиться в одном из следующих состояний:

E$_0$: k = 0 – в системе нет заявок (оба прибора простаивают);

E$_1$: k = 1 – в системе (на обслуживании в одном из приборов) находится 1 заявка;

E$_2$: k = 2 – в системе (на обслуживании в обоих приборах) находятся

2 заявки;

E$_3$: k = 3 – в системе находятся 3 заявки: две – на обслуживании в приборах и одна – в накопителе.

4. Размеченный граф переходов случайного процесса представлен на рис.5.13.

В один и тот же момент времени может произойти одно из двух событий, которые приводят к изменению состояния случайного процесса, протекающего в системе:

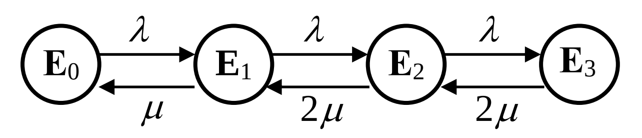


Рис.5.13. Размеченный граф переходов

• поступление заявки в систему с интенсивностью λ, приводящее к переходу в соседнее состояние с бóльшим номером, причем если случайный процесс находится в состоянии **E**$_3$, то его состояние не изменится, что соответствует отказу в обслуживании поступившей заявке;

• завершение обслуживания заявки в одном из приборов с интенсивностью μ, при этом случайный процесс переходит в соседнее состояние с меньшим номером с интенсивностью μ, если работает один прибор, и с интенсивностью 2μ, если работают оба прибора.

5. Расчет характеристик СМО.

Не выписывая матрицу интенсивностей переходов и систему уравнений для определения стационарных вероятностей, запишем формулы для расчёта характеристик СМО:

1) нагрузка: $y = \lambda / \mu = \lambda b$;

2) загрузка: $\rho = (p_1 + 2p_2 + 2p_3)/2$;

3) среднее число работающих приборов: $k' = (p_1 + 2p_2 + 2p_3) = 2\rho$;

4) коэффициент простоя системы: $\eta = 1 - \rho$;

5) среднее число заявок в очереди: $l = p_3$;

6) среднее число заявок в системе: $m = p_1 + 2p_2 + 3p_3 = l + k'$;

7) вероятность потери заявок: $\pi = p_3$;

Задание на самостоятельную работу: *используя выражение (3.18), доказать, что вероятность потери заявок равна вероятности того, что система заполнена, то есть в накопителе нет свободных мест для вновь поступающих заявок.*

8) производительность системы (интенсивность потока обслуженных заявок): $\lambda' = \lambda(1 - \pi)$;

9) интенсивность потока потерянных заявок: $\lambda'' = \lambda\pi$;

10) среднее время ожидания заявок: $w = l / \lambda'$;

11) среднее время пребывания заявок: $u = m / \lambda' = w + b$.

5.4.6. Одноканальная СМО с неоднородным потоком

заявок и относительными приоритетами

Рассмотрим одноканальную СМО, в которую поступает неоднородный поток заявок. Ожидающие обслуживания заявки разнесены по разным накопителям ограниченной ёмкости. Между заявками разных классов установлены относительные приоритеты (ОП), означающие, что всякий раз из накопителей на обслуживание выбирается заявка с самым высоким приоритетом. При этом при поступлении в систему высокоприоритетной заявки обслуживание низкоприоритетной не прерывается. При заполненных накопителях поступившая заявка теряется.

1. Описание системы (рис.5.14).

1.1. Система *одноканальная*.

1.2. Входящий поток заявок – *неоднородный:* в систему поступает два класса заявок.

1.3. Накопители для заявок каждого класса – *ограниченной ёмкости*: $r_1 = r_2 = 1$.

1.4. Дисциплина буферизации – *без вытеснения* заявок: если при поступлении в систему заявки любого класса соответствующий накопитель заполнен до конца, то заявка теряется

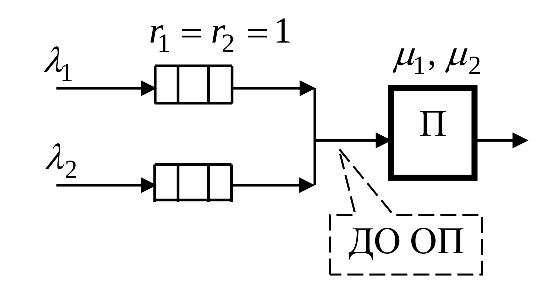


Рис.5.14. СМО с неоднородным потоком заявок

1.4. Дисциплина обслуживания – *с относительными приоритетами*: заявки первого класса имеют приоритет по отношению к заявкам второго класса.

2. Предположения и допущения.

2.1. Поступающие в систему заявки двух классов образуют *простейшие* потоки с интенсивностями λ_1 и λ_2 соответственно.

2.2. Длительности обслуживания заявок каждого класса распределены по *экспоненциальному* закону с интенсивностями $\mu_1 = 1/b_1$ и $\mu_2 = 1/b_2$, где b_1 и b_2 – средние длительности обслуживания заявок класса 1 и 2 соответственно.

В СМО всегда существует стационарный режим, так как не может быть бесконечных очередей.

3. Кодирование состояний случайного процесса.

Для описания состояний марковского процесса будем использовать распределение заявок между прибором и накопителями. Закодируем состояния следующим образом: (П/O_1,O_2), где П = {0, 1, 2} – состояние обслуживающего прибора, задаваемое классом заявки, находящейся на обслуживании («0» – прибор свободен; «1» или «2» – на обслуживании в приборе находится заявка класса 1 или 2 соответственно); O_1, O_2 = {0, 1} – состояние накопителей 1 и 2 соответственно («0» – означает отсутствие заявки в накопителе, «1» –означает наличие одной заявки в накопителе соответствующего класса).

При выбранном способе кодирования система может находиться в следующих состояниях:

E$_0$: (0/0,0) – в системе нет ни одной заявки;

E$_1$: (1/0,0) – на обслуживании в приборе находится заявка класса 1;

E$_2$: (2/0,0) – на обслуживании в приборе находится заявка класса 2;

E$_3$: (1/1,0) – на обслуживании находится заявка класса 1 и одна заявка класса 1 ожидает обслуживания в первом накопителе;

E$_4$: (1/0,1) – на обслуживании находится заявка класса 1 и одна заявка класса 2 ожидает обслуживания соответственно во втором накопителе;

E$_5$: (2/1,0) – на обслуживании находится заявка класса 2 и одна заявка класса 1 ожидает обслуживания в первом накопителе;

E$_6$: (2/0,1) – на обслуживании находится заявка класса 2 и одна заявка класса 2 ожидает обслуживания во втором накопителе;

E$_7$: (1/1,1) – на обслуживании находится заявка класса 1, и по одной заявке каждого класса ожидают обслуживания в соответствующих накопителях;

E$_8$: (2/1,1) – на обслуживании находится заявка класса 2, и по одной заявке каждого класса ожидают обслуживания в соответствующих накопителях.

Отметим, что при кодировании случайных процессов могут быть применены различные способы кодирования.

В рассматриваемом примере состояния случайного процесса вместо представленного выше способа можно закодировать, например, следующим образом: (П,О), где П = {0, 1, 2} – состояние обслуживающего прибора, задаваемое классом заявки, находящейся на обслуживании («0» – прибор свободен; «1» или «2» – на обслуживании в приборе находится заявка класса 1 или 2 соответственно); О = {0, 1, 2, 3} – состояние накопителей 1 и 2 соответственно («0» – означает отсутствие заявок в обоих накопителях; «1» –наличие в первом накопителе заявки класса 1; «2» – наличие во втором накопителе заявки класса 2; «3» – наличие в первом и втором накопителях по одной заявке соответственно класса 1 и 2).

Представленные способы кодирования не применимы, если для заявок обоих классов используется общий накопитель ёмкостью $r = 2$. В этом случае количество состояний случайного процесса увеличится, поскольку в накопителе могут находиться 2 заявки одного и того же класса и состояние накопителя может быть представлено следующим образом: О = {0, 1, 2, 11, 12, 22}, где «0» – означает отсутствие заявок в накопителе; «1» – наличие в накопителе только одной заявки класса 1; «2» – наличие в накопителе заявки класса 2; «11» – наличие в накопителе двух заявок класса 1; «22» – наличие в накопителе двух заявок класса 2 и «12» – наличие в накопителе одной заявки класса 1 и одной заявки класса 2. Заметим, что состояние «12» не различает, в какой последовательности эти заявки поступили в систему, что обусловлено наличием относительного приоритета между ними – независимо от момента поступления на обслуживание первой всегда будет выбрана заявка класса 1. В случае бесприоритетного обслуживания, когда заявки разных классов выбираются на обслуживание в порядке поступления, следует ввести ещё одно состояние накопителя – «21», означающее, что заявка класса 2 поступила в систему раньше заявки класса 1, в то время как состояние «12» означает, что в систему раньше поступила заявка класса 1.

4. Размеченный граф переходов случайного процесса представлен на рис.5.15.

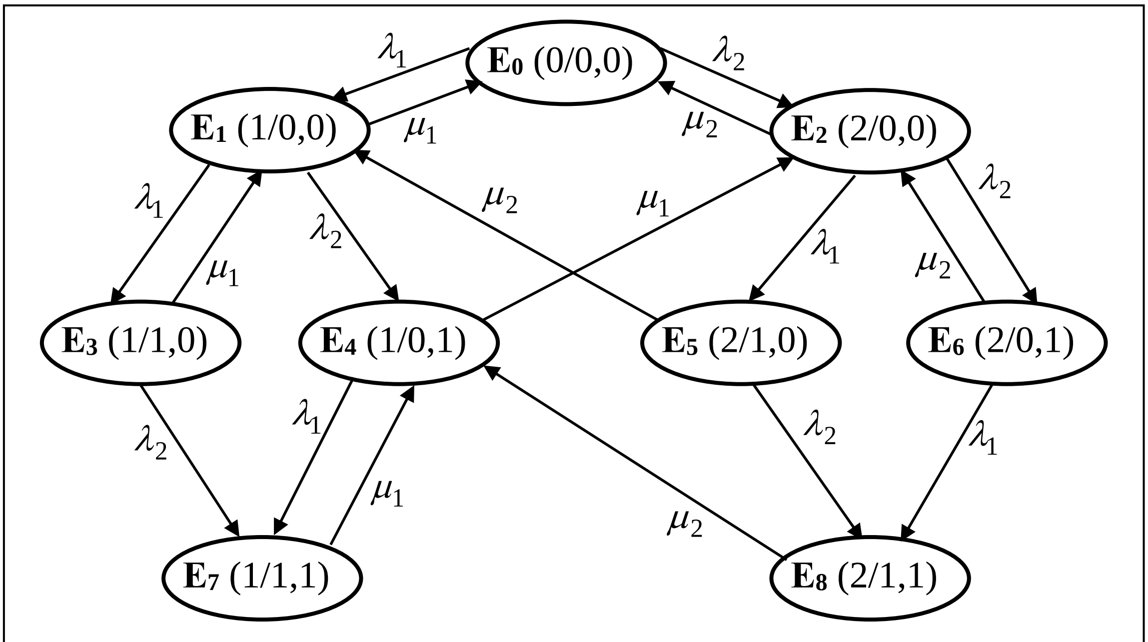


Рис.5.15. Размеченный граф переходов

В каждый момент времени может произойти только одно событие (или поступление заявки какого-либо класса, или завершение обслуживания заявки, находящейся в приборе), поскольку вероятность появления двух и более событий в один и тот же момент времени равна нулю.

При наличии в накопителях заявок первого и второго классов (состояния $\mathbf{E}_7$ и $\mathbf{E}_8$) после завершения обслуживания некоторой заявки в приборе случайный процесс переходит в состояние $\mathbf{E}_4$, означающее, что на обслуживание всегда выбирается высокоприоритетная заявка класса 1.

По графу переходов составим систему уравнений для определения стационарных вероятностей:

$$\begin{cases}(\lambda_1+\lambda_2)\,p_0=\mu_1 p_1+\mu_2 p_2\\(\lambda_1+\lambda_2+\mu_1)\,p_1=\lambda_1 p_0+\mu_1 p_3+\mu_2 p_5\\(\lambda_1+\lambda_2+\mu_2)\,p_2=\lambda_2 p_0+\mu_1 p_4+\mu_2 p_6\\(\lambda_2+\mu_1)\,p_3=\lambda_1 p_1\\(\lambda_1+\mu_1)\,p_4=\lambda_2 p_1+\mu_1 p_7+\mu_2 p_8\\(\lambda_2+\mu_2)\,p_5=\lambda_1 p_2\\(\lambda_1+\mu_2)\,p_6=\lambda_2 p_2\\\mu_1 p_7=\lambda_2 p_8+\lambda_1 p_4\\\mu_2\,p_8=\lambda_2 p_5+\lambda_1 p_6\\p_0+p_1+p_2+p_3+p_4+p_5+p_6+p_7+p_8=1\end{cases}$$

5. Расчет характеристик СМО.

Характеристики обслуживания заявок в СМО с неоднородным потоком заявок делятся на две группы:

- характеристики обслуживания заявок каждого класса;
- характеристики обслуживания заявок суммарного потока.

Расчёт характеристик обслуживания заявок *каждого класса* выполняется по следующим формулам:

1) нагрузка: $y_1=\lambda_1/\mu_1=\lambda_1 b_1$; $y_2=\lambda_2/\mu_2=\lambda_2 b_2$;

2) загрузка, создаваемая заявками, которая может трактоваться как вероятность того, что на обслуживании в приборе находится заявка класса 1 и 2 соответственно: $\rho_1=p_1+p_3+p_4+p_7$; $\rho_2=p_2+p_5+p_6+\rho_8$;

3) среднее число заявок в очереди:

$l_1=p_3+p_5+p_7+p_8$; $l_2=p_4+p_6+p_7+p_8$;

4) среднее число заявок в системе:

$m_1=p_1+2p_3+p_4+p_5+2p_7+p_8=l_1+\rho_1$;

$m_2=p_2+p_4+p_5+2p_6+p_7+2p_8=l_2+\rho_2$;

5) вероятность потери заявок:

$\pi_1=p_3+p_5+p_7+p_8$; $\pi_2=p_4+p_6+p_7+p_8$;

6) производительность по каждому классу заявок (интенсивность непотерянных заявок):

$\lambda_1^{'}=\lambda_1(1-\pi_1)$; $\lambda_2^{'}=\lambda_2(1-\pi_2)$;

7) среднее время ожидания заявок:

$w_1=l_1/\lambda_1^{'}$; $w_2=l_2/\lambda_2^{'}$

8) среднее время пребывания заявок

$u_1=m_1/\lambda_1^{'}=w_1+b$; $u_2=m_2/\lambda_2^{'}=w_2+b$

Расчёт характеристик обслуживания заявок *суммарного потока* выполняется по следующим формулам:

1) суммарная нагрузка системы: $Y=y_1+y_2$;

2) загрузка системы: $R = \rho_1 + \rho_2$;

3) коэффициент простоя системы: $\eta = 1 - R$;

4) суммарное число заявок во всех очередях: $l = l_1 + l_2$;

5) суммарное число заявок в системе: $m = m_1 + m_2 = l + R$;

6) вероятность потери заявок: $\pi = \pi_1 + \pi_2$;

7) производительность системы (интенсивность суммарного потока обслуженных заявок): $\lambda' = \lambda_1' + \lambda_2' = \lambda(1 - \pi)$;

8) среднее время ожидания заявок суммарного потока:
$w = (\lambda_1' w_1 + \lambda_2' w_2) / \lambda' = l / \lambda'$;

9) среднее время пребывания заявок суммарного потока:
$u = (\lambda_1' u_1 + \lambda_2' u_2) / \lambda' = m / \lambda' = w + b$.

5.5. Марковские модели сетей массового обслуживания

В данном параграфе подробно рассматриваются марковские модели сетей массового обслуживания (СеМО) с однородным потоком заявок. В качестве примеров представлены разомкнутые и замкнутые экспоненциальные СеМО с накопителями ограниченной ёмкости, а также замкнутые неэкспоненциальные СеМО, в которых длительность обслуживания заявок в одном из узлов распределена по закону Эрланга с коэффициентом вариации $\nu < 1$ и гиперэкспоненциальному закону с коэффициентом вариации $\nu > 1$.

Можно показать, что случайный процесс, протекающий в экспоненциальных разомкнутых и замкнутых СеМО при сформулированных предположениях и допущениях является марковским.

Случайный процесс, протекающий в замкнутой неэкспоненциальной сети, не является марковским. Для описания процесса функционирования такой системы в терминах марковских случайных процессов в некоторых случаях можно воспользоваться методом вложенных цепей Маркова, суть которого заключается в том, что функционирование системы рассматривается в определенные моменты времени, образующие цепь Маркова.

Как и для СМО, в каждом примере приводится *описание* исследуемой СеМО и принятые при построении математической модели *предположения и допущения*, необходимые для того, чтобы протекающий в системе случайный процесс мог быть сведён к марковскому. Разработка Марковской модели включает в себя этапы *кодирования состояний* случайного процесса, построения размеченного *графа переходов*, формирования *матрицы интенсивностей переходов* и *системы линейных алгебраических уравнений* для расчёта стационарных вероятностей состояний марковского процесса, на основе которых строятся математические зависимости, позволяющие рассчитать наиболее важные характеристики функционирования исследуемых СеМО.

Применение марковских случайных процессов для расчёта характеристик функционирования и исследования свойств СеМО оказывается наиболее результативным:

- для разомкнутых СеМО с накопителями ограниченной ёмкости, в которых заявки теряются при заполненных накопителях;
- для неэкспоненциальных разомкнутых и замкнутых СеМО, в которых длительности обслуживания заявок в узлах распределены по гипоэкспоненциальному или гиперэкспоненциальному закону.

5.5.1. Разомкнутая экспоненциальная СеМО с накопителями ограниченной емкости

Рассмотрим разомкнутую экспоненциальную СеМО с двумя одноканальными узлами, в которую из внешней среды поступает простейший поток заявок с интенсивностью λ_0 (рис.5.16). Накопители в обоих узлах имеют ограниченную ёмкость, равную единице: $r_1 = r_2 = 1$. Заявка, поступившая в узел и заставшая накопитель заполненным, теряется. Длительности обслуживания в узлах распределены по экспоненциальному закону со средними значениями b_1 и b_2 соответственно. Заявки после обслуживания в узле 2 вероятностью q направляются в узел 1 и вероятностью $(1-q)$ – покидают СеМО.

Отметим, что, поскольку заявки в сети могут теряться, рассматриваемая разомкнутая СеМО является *нелинейной*, то есть интенсивности потоков заявок, поступающих в узлы СеМО, не связаны между собой линейной зависимостью (3.5) и, следовательно, не могут быть рассчитаны путём решения системы линейных алгебраических уравнений (4.16).

1. Описание СеМО (рис.5.16)**.**

1.1. Сеть массового обслуживания – разомкнутая *двухузловая*.

1.2. Узлы 1 и 2 – одноканальные: $K_1 = K_2 = 1$.

1.3. Накопители в узлах ограниченной ёмкости: $r_1 = r_2 = 1$.

1.4. Дисциплины буферизации в узлах – с потерями заявок, если накопители заполнены.

1.5. Поток заявок *однородный с интенсивностью* λ_0.

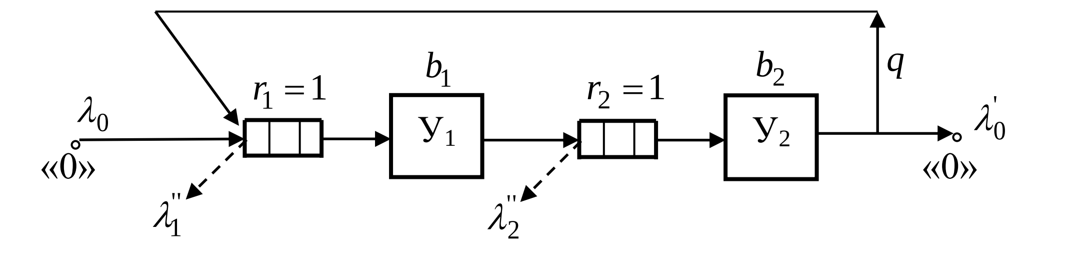


Рис. 5.16. Двухузловая разомкнутая СеМО с потерями

2. Предположения и допущения.

2.1. Поступающие в разомкнутую СеМО заявки образуют *простейший* поток с интенсивностью λ_0.

2.2. Длительности обслуживания заявок в узлах СеМО распределены по *экспоненциальному* закону с параметрами, представляющими собой интенсивности обслуживания: $\mu_1 = 1/b_1$ и $\mu_2 = 1/b_2$.

В разомкнутой СеМО при любой нагрузке существует стационарный режим, так как в узлах сети не может быть бесконечных очередей.

3. Кодирование состояний случайного процесса.

Для описания состояний марковского случайного процесса будем использовать распределение заявок между узлами. Закодируем состояния следующим образом: (M_1, M_2), где M_i = {0, 1, 2} – количество заявок в узле i («0» – узел свободен; «1» – на обслуживании в узле находится одна заявка; «2» – в узле находятся две заявки – одна на обслуживания и вторая в накопителе).

При выбранном способе кодирования система может находиться в следующих состояниях:

E$_0$: (0,0) – в СеМО нет ни одной заявки;

E$_1$: (1,0) – в узле 1 находится одна заявка;

E$_2$: (2,0) – в узле 1 находятся две заявки;

E$_3$: (0,1) – в узле 2 находится одна заявка;

E$_4$: (1,1) – в узле 1 и 2 находится по одной заявке;

E$_5$: (2,1) – две заявки находятся в узле 1 и одна – в узле 2;

E$_6$: (0,2) – в узле 2 находятся две заявки;

E$_7$: (1,2) – две заявки находятся в узле 2 и одна – в узле 1;

E$_8$: (2,2) – в узле 1 и 2 находятся по две заявки.

4. Размеченный граф переходов случайного процесса (рис.5.17).

Построим граф переходов, полагая, что в каждый момент времени может произойти только одно событие (поступление заявки в СеМО или завершение обслуживания заявки в одном из узлов), поскольку вероятность появления двух и более событий в один и тот же момент времени равна нулю.

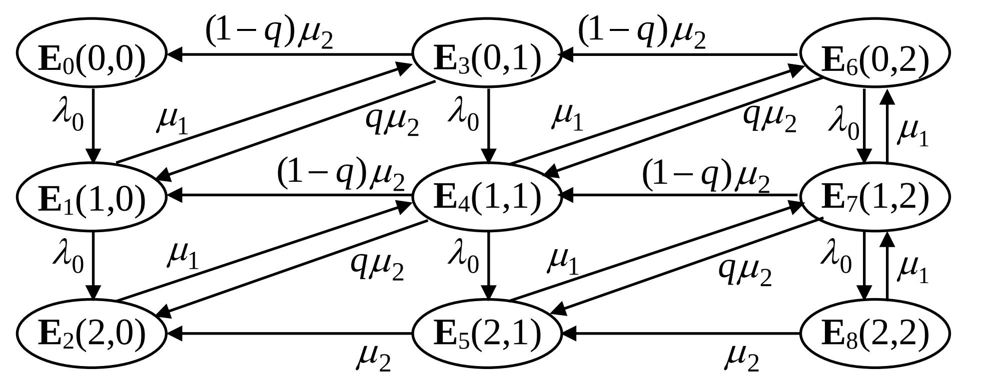


Рис. 5.17. Граф переходов разомкнутой СеМО с потерями

Следует обратить внимание на переходы из состояний **E**$_3$(0,1), **E**$_4$(1,1), **E**$_6$(0,2) и **E**$_7$(1,2), обусловленные завершением обслуживания заявки в узле 2 с интенсивностью μ_2. В этих случаях с вероятностью q заявка может вернуться в узел 1 и с вероятностью $(1-q)$ – покинуть СеМО, тогда интенсивности соответствующих переходов будут равны $q\mu_2$ и $(1-q)\mu_2$. Если же случайный процесс находится в состояниях **E**$_5$(2,1) и **E**$_8$(2,2), то завершение обслуживания заявки в узле 2 приводит к переходу соответственно в состояния **E**$_2$(2,0) и **E**$_5$(2,1) с интенсивностью μ_2, что соответствует выходу заявки из СеМО с вероятностью $(1-q)$ и потере заявки, которая с вероятностью q будет направлена в узел 1, поскольку в последнем нет места в накопителе. Аналогично, если случайный процесс находится в состояниях **E**$_7$(1,2) и **E**$_8$(2,2), то завершение обслуживания заявки в узле 1 приводит к переходу соответственно в состояния **E**$_6$(0,2) и **E**$_7$(1,2) с интенсивностью μ_1, что соответствует потере заявки, поскольку накопитель узла 1 заполнен.

5. Расчет характеристик СеМО.

Не составляя матрицу интенсивностей переходов и не выписывая систему уравнений для определения стационарных вероятностей, получим математические выражения для определения узловых и сетевых характеристик разомкнутой СеМО с потерями при известных значениях стационарных вероятностей состояний $p_i \quad (i = 0,1,\ldots,8)$.

Заметим, что СеМО с потерями относится к классу *нелинейных* сетевых моделей, расчёт характеристик которых связан с определёнными проблемами, в частности, с необходимостью детального анализа потоков заявок и с невозможностью применения в ряде случаев фундаментальных соотношений для расчёта сетевых характеристик. Кроме того, процесс формирования математических зависимостей для каждой конкретной нелинейной СеМО может существенно отличаться.

В связи с этим, ниже достаточно подробно рассматривается процесс получения математических выражений для расчёта узловых и сетевых характеристик нелинейной разомкнутой СеМО, представленной на рис.5.16.

Узловые характеристики СеМО рассчитываются в такой последовательности:

1) загрузки узлов определяются как суммы вероятностей состояний, в которых соответствующий узел занят обслуживанием заявок:

$\rho_1 = p_1 + p_2 + p_4 + p_5 + p_7 + p_8$; $\rho_2 = p_3 + p_4 + p_5 + p_6 + p_7 + p_8$;

2) коэффициенты простоя узлов: $\eta_1 = 1-\rho_1; \quad \eta_2 = 1-\rho_2$;

3) среднее число заявок в очередях:

$l_1 = p_2 + p_5 + p_8; \qquad l_2 = p_6 + p_7 + p_8$;

4) среднее число заявок в узлах:

$m_1 = p_1 + p_4 + p_7 + 2(p_2 + p_5 + p_8) = l_1 + \rho_1$;

$m_2 = p_3 + p_4 + p_5 + 2(p_6 + p_7 + p_8) = l_2 + \rho_2$;

5) производительности узлов (интенсивность обслуженных заявок на

выходе узлов): $\lambda_1^{'} = \frac{\rho_1}{b_1} = \rho_1\mu_1$; $\lambda_2^{'} = \frac{\rho_2}{b_2} = \rho_2\mu_2$;

6) вероятности потери заявок в узлах СеМО могут быть рассчитаны на основе выражения (3.18) с учётом того, что $K_1 = K_2 = 1$:

$$\pi_1 = 1 - \frac{\rho_1}{y_1}; \qquad \pi_2 = 1 - \frac{\rho_2}{y_2};$$

в этих выражениях: $y_1 = \lambda_1 b_1$ и $y_2 = \lambda_2 b_2$ – создаваемые в узлах нагрузки, где λ_1 и λ_2 – интенсивности поступления заявок в узлы 1 и 2 СеМО, для расчёта которых необходимо выполнить анализ потоков в рассматриваемой СеМО;

интенсивность λ_1 складывается (см. рис.5.16) из интенсивности λ_0 поступления заявок из внешнего источника и интенсивности потока заявок, возвращающихся с вероятностью q в узел 1 после обслуживания в узле 2: $\lambda_1 = \lambda_0 + q\lambda_2^{'}$, где $\lambda_2^{'}$ - рассчитанная ранее интенсивность потока выходящих из узла 2 заявок (производительность узла 2);

аналогично, из рис. 5.16 можно видеть, что интенсивность λ_2 поступающих в узел 2 заявок представляет собой интенсивность $\lambda_1^{'}$ потока выходящих из узла 1 заявок (производительность узла 1): $\lambda_2 = \lambda_1^{'}$;

окончательно, после некоторых преобразований выражения для расчёта вероятностей потери заявок в узлах СеМО примут вид:

$$\pi_1 = 1 - \frac{\lambda_1^{'}}{\lambda_0 + q\lambda_2^{'}}; \qquad \pi_2 = 1 - \frac{\lambda_2^{'}}{\lambda_1^{'}};$$

7) среднее время ожидания заявок в узлах рассчитывается по формулам Литтла с учётом только обслуженных заявок:

$w_1 = l_1 / \lambda_1^{'}$; $w_2 = l_2 / \lambda_2^{'}$;

8) аналогично, среднее время пребывания заявок в узлах:

$u_1 = m_1 / \lambda_1^{'} = w_1 + b$; $u_2 = m_2 / \lambda_2^{'} = w_2 + b$;

Для расчёта *сетевых характеристик* СеМО могут использоваться следующие формулы:

1) суммарная загрузка узлов СеМО, характеризующая среднее число одновременно работающих узлов в сети: $\rho = \rho_1 + \rho_2$;

2) суммарное число заявок в очередях: $L = l_1 + l_2$;

3) суммарное число заявок в узлах: $M = m_1 + m_2 = L + \rho$;

4) производительность СеМО (интенсивность обслуженных заявок на выходе сети): $\lambda_0^{'} = (1-q)\lambda_2^{'}$;

5) вероятность потери заявок в сети: $\pi = \frac{\lambda_0 - \lambda_0^{'}}{\lambda_0} = 1 - \frac{\lambda_0^{'}}{\lambda_0}$; следует обратить внимание, что вероятность потери заявок в сети определяется как

доля потерянных заявок по отношению к поступившим в СеМО заявкам, в то время как вероятности потери π_1 и π_2 заявок в узлах СеМО определяется как доля потерянных заявок по отношению ко всем заявкам, поступившим в конкретный узел, число которых учитывает и то, что поступившая в СеМО заявка за время нахождения в сети может попасть в данный узел несколько раз.

Математические зависимости для расчёта суммарного времени ожидания заявок и времени пребывания заявок в СеМО не могут быть получены в общем виде в виду нелинейности СеМО с потерями.

5.5.2. Замкнутая экспоненциальная СеМО

1. Описание замкнутой СеМО (рис.5.18).

1.1. Сеть массового обслуживания (СеМО) – замкнутая *двухузловая*.

1.2. Количество приборов в узлах: узел 1 – одноканальный, узел 2 – двухканальный.

1.3. Поток заявок *однородный*.

1.4. В СеМО постоянно циркулируют $M = 3$ заявки.

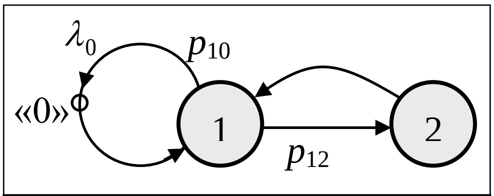


Рис.5.18. Замкнутая СеМО

2. Предположения и допущения.

2.1. Длительности обслуживания заявок в узлах 1 и 2 распределены по экспоненциальному закону с интенсивностями $\mu_1 = 1/b_1$ и $\mu_2 = 1/b_2$ соответственно, где b_1, b_2 – средние длительности обслуживания заявок.

2.2. Приборы в двухканальном узле 2 *идентичны* и любая заявка может обслуживаться в любом приборе.

2.3. Заявка после обслуживания в узле 1 с вероятностью p_{12} переходит в узел 2 и с вероятностью $p_{10} = 1 - p_{12}$ возвращается в этот же узел 1.

2.4. Дуга, выходящая из узла 1 и входящая обратно в этот же узел, рассматривается как внешняя по отношению к СеМО, и на ней выбирается нулевая точка «0».

В замкнутой СеМО всегда существует стационарный режим, так как число заявок в сети ограничено и не может быть бесконечных очередей.

Легко убедиться, что случайный процесс, протекающий в замкнутой экспоненциальной сети, является марковским.

3. Кодирование состояний марковского процесса.

Под состоянием марковского процесса будем понимать распределение заявок по узлам СеМО. Закодируем состояния следующим образом: (**M₁, M₂**), где **M₁** = {0, 1, 2, 3} – количество заявок, находящихся в узле 1 и **M₂** = {0, 1, 2, 3} – количество заявок, находящихся в узле 2, причем суммарное число заявок в обоих узлах должно быть равно 3.

При выбранном способе кодирования система может находиться в следующих состояниях:

E₀: (3, 0) – все три заявки находятся в узле 1, причем одна заявка находятся на обслуживании в приборе и две заявки ожидают в накопителе;

E$_1$: (2, 1) – две заявки находятся в узле 1 (одна на обслуживании в приборе и одна в накопителе) и одна – на обслуживании в одном из приборов узла 2;

E$_2$: (1, 2) – одна заявка находится на обслуживании в узле 1 и две – в узле 2 (на обслуживании в обоих приборах);

E$_3$: (0, 3) – все три заявки находятся в узле 2, причем две заявки находятся на обслуживании в обоих приборах узла 2 и одна заявка ожидает в накопителе.

4. Размеченный граф переходов случайного процесса (рис.5.19).

В один и тот же момент времени в замкнутой СеМО может произойти только одно из двух событий:

1) завершение обслуживания заявки в первом узле с интенсивностью μ_1, при этом заявка с вероятностью p_{12} покинет этот узел и перейдет в узел 2 (интенсивность перехода $p_{12}\mu_1$) или с вероятностью $(1-p_{12})$ останется в этом же узле, то есть состояние случайного процесса не изменится; отметим, что второй случай не отображается на графе переходов в виде дуги, выходящей из узла 1 и снова входящей в узел 1;

2) завершение обслуживания заявки в узле 2 с интенсивностью μ_2, если на обслуживании в этом узле находится одна заявка (работает один прибор), или с интенсивностью $2\mu_2$, если на обслуживании в узле находятся две заявки (работают оба прибора); обслуженная заявка покидает этот узел и с вероятностью 1 переходит в первый узел.

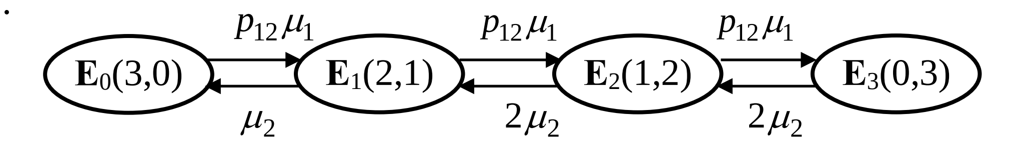


Рис.5.19. Граф экспоненциальной ЗСеМО

5. Система уравнений.

Не составляя матрицу интенсивностей переходов, запишем систему уравнений для определения стационарных вероятностей:

$$\begin{cases} p_{12}\mu_1 p_0 = \mu_2 p_1 \\ (p_{12}\mu_1 + \mu_2) p_1 = p_{12}\mu_1 p_0 + 2\mu_2 p_2 \\ (p_{12}\mu_1 + 2\mu_2) p_2 = p_{12}\mu_1 p_1 + 2\mu_2 p_3 \\ 2\mu_2 p_3 = p_{12}\mu_1 p_2 \\ p_0 + p_1 + p_2 + p_3 = 1 \end{cases}$$

6. Расчет характеристик СеМО.

На основе полученных значений стационарных вероятностей рассчитываются узловые и сетевые характеристики СеМО с использованием следующих формул:

1) загрузка узлов:

$\rho_1 = p_0 + p_1 + p_2; \quad \rho_2 = 0{,}5p_1 + p_2 + p_3;$

2) коэффициенты простоя узлов:

$\eta_1 = 1 - \rho_1; \quad \eta_2 = 1 - \rho_2;$

3) средние длины очередей заявок в узлах:

$l_1 = 2p_0 + p_1; \quad l_2 = p_3;$

4) среднее число заявок в узлах:

$m_1 = 3p_0 + 2p_1 + p_2; \quad m_2 = p_1 + 2p_2 + 3p_3;$

5) производительность замкнутой СеМО:

$$\lambda_0 = \frac{\rho_1}{\alpha_1 b_1} = \frac{\rho_2}{\alpha_2 b_2};$$

где α_1 и α_2 – коэффициенты передач соответственно узлов 1 и 2, определяемые путем решения системы уравнений (4.17);

6) среднее время ожидания заявок в узлах СеМО:

$$w_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad w_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

7) среднее время пребывания заявок в узлах СеМО:

$$u_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad u_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

8) нагрузка в узлах сети:

$y_1 = \alpha_1 \lambda_0 b_1; \quad y_2 = \alpha_2 \lambda_0 b_2;$

9) среднее число параллельно работающих *узлов* сети, определяемое как суммарная *загрузка* всех узлов СеМО:

$R = \rho_1 + \rho_2;$

10) среднее число параллельно работающих *приборов* во всех узлах сети, определяемое как суммарная *нагрузка* всех узлов СеМО:

$Y = y_1 + y_2;$

11) суммарное число заявок во всех очередях СеМО:

$L = l_1 + l_2;$

12) суммарное (полное) время ожидания заявок в СеМО :

$W = \alpha_1 w_1 + \alpha_2 w_2;$

13) время пребывания заявок в СеМО:

$U = \alpha_1 u_1 + \alpha_2 u_2;$

Суммарное число заявок, циркулирующих в ЗСеМО, рассчитываемое как $M = m_1 + m_2$, должно совпадать с заданным числом заявок $M = 3$.

Следует обратить внимание на то, что временные характеристики обслуживания заявок в узлах СеМО, а, следовательно, и в сети в целом, могут быть рассчитаны только после определения производительности замкнутой СеМО, вычисляемой через найденные значения загрузок узлов.

5.5.3. Замкнутая СеМО с эрланговским обслуживанием

1. Описание СеМО.

1.1. Замкнутая сеть массового обслуживания (ЗСеМО) – *двухузловая*.

1.2. Количество приборов в узлах: $K_1 = K_2 = 1$.

1.3. Поток заявок *однородный*.

1.4. В ЗСеМО постоянно циркулируют $M = 3$ заявки.

Граф рассматриваемой ЗСеМО такой же, как и в предыдущем примере (рис.5.18). Отличие состоит только в том, что в рассматриваемой ЗСеМО узел 2 – одноканальный.

2. Предположения и допущения.

2.1. Длительность обслуживания заявок *в узле 1* распределена по *закону Эрланга 2-го порядка* со средней длительностью обслуживания заявок $b_1 = 1/\mu_1$, а в узле 2 – по экспоненциальному закону со средней длительностью обслуживания заявок $b_2 = 1/\mu_2$, где μ_1, μ_2 – интенсивности обслуживания заявок.

2.2. Заявка после обслуживания в узле 1 с вероятностью p_{12} переходит в узел 2 и с вероятностью $p_{10} = 1 - p_{12}$ возвращается в этот же узел 1.

2.3. Дуга, выходящая из узла 1 и входящая обратно в этот же узел, рассматривается как внешняя по отношению к СеМО, и на ней отмечается нулевая точка «0».

3. Сведение случайного процесса к марковскому.

Случайный процесс, протекающий в замкнутой неэкспоненциальной сети, не является марковским.

Для описания процесса функционирования такой системы в терминах марковских случайных процессов будем рассматривать функционирование системы в определенные моменты времени, в которые случайный процесс обладает марковским свойством. Для этого воспользуемся представлением случайной величины, распределенной по закону Эрланга 2-го порядка, в виде суммы двух экспоненциально распределенных случайных величин (см. раздел 2, п.2.6.1). При этом будем полагать, что обслуживание заявки в первом узле проходит две фазы, длительность каждой из которых распределена по экспоненциальному закону со средним значением $b_1^{'} = b_1 / 2$. Последнее необходимо для того, чтобы полная длительность обслуживания в узле 1 была равна b_1 .

Таким образом, обслуживание заявки в СеМО можно представить как двухфазное обслуживание в первом узле и однофазное – во втором узле (рис.5.20). Длительности обслуживания в фазах Ф1 и Ф2 первого узла ЗСеМО распределены по экспоненциальному закону с одним и тем же параметром $\mu_1^{'} = 1/b_1^{'}$ и с параметром $\mu_2 = 1/b_2$ – в единственной фазе второго узла. Моменты завершения обслуживания в каждой из фаз образуют цепь Маркова, так как времена нахождения в них распределены по экспоненциальному закону. Такое представление случайного процесса требует другого подхода к кодированию состояний, учитывающего распределение заявок по фазам обслуживания.

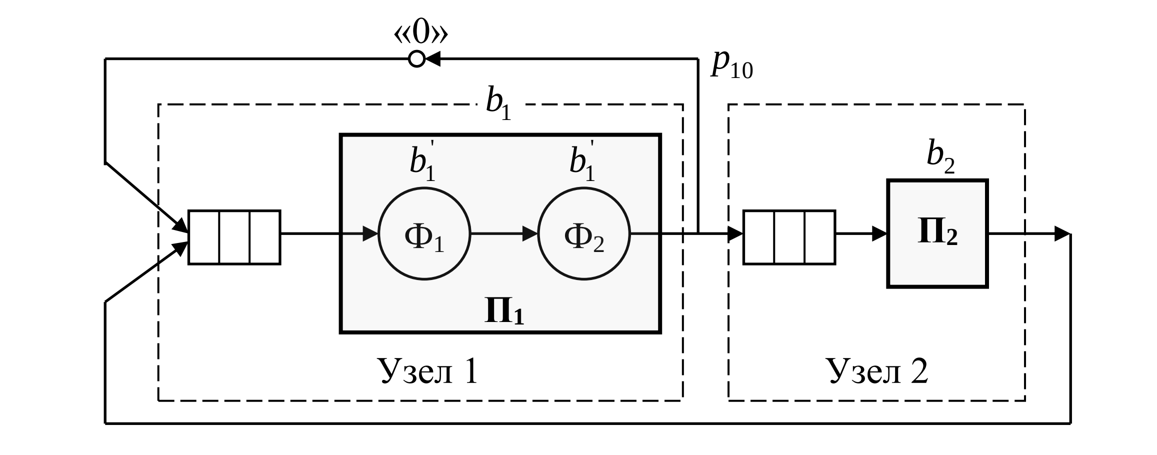


Рис.5.20. ЗСеМО с двухфазным представлением распределения Эрланга

4. Кодирование состояний случайного процесса.

Под состоянием марковского процесса будем понимать распределение заявок по узлам СеМО с учетом того, на какой фазе обслуживания распределения Эрланга находится заявка в узле 1.

Для этого закодируем состояния следующим образом: (**M₁, M₂**), где **M₁** = {0, 1_1, 1_2, 2_1, 2_2, 3} – количество заявок, находящихся в узле 1 (индексы отражают нахождение заявки на 1-й или 2-й фазе распределения Эрланга), и **M₂** = {0, 1, 2, 3} – количество заявок, находящихся в узле 2, причем суммарное число заявок в обоих узлах должно быть равно 3.

При выбранном способе кодирования система может находиться в следующих состояниях:

E$_1$: (3_1, 0) – все три заявки находятся в узле 1, причем одна заявка находятся на обслуживании в приборе на *первой фазе*, и две заявки ожидают в накопителе;

E$_2$: (3_2, 0) – все три заявки находятся в узле 1, причем одна заявка находятся на обслуживании в приборе на *второй фазе*, и две заявки ожидают в накопителе;

E$_3$: (2_1, 1) – две заявки находятся в узле 1 (одна на обслуживании в приборе на *первой фазе* и одна в накопителе) и одна – на обслуживании в узле 2;

E$_4$: (2_2, 1) – две заявки находятся в узле 1 (одна на обслуживании в приборе на *второй фазе* и одна в накопителе) и одна – на обслуживании в узле 2;

E$_5$: (1_1, 2) – одна заявка находится в узле 1 на обслуживании в приборе на *первой фазе* и две заявки находятся в узле 2, причем одна из них находится на обслуживании в приборе, а вторая заявка ожидает в накопителе;

E$_6$: (1_2, 2) – одна заявка находится в узле 1 на обслуживании в приборе на *второй фазе* и две заявки находятся в узле 2, причем одна из них находится на обслуживании в приборе, а вторая заявка ожидает в накопителе;

E$_7$: (0, 3) – все три заявки находятся в узле 2, причем одна заявка находится на обслуживании в приборе, а две другие – ожидают в накопителе.

5. Размеченный граф переходов случайного процесса.

На рис.5.21 представлен граф переходов марковского процесса для рассматриваемой неэкспоненциальной СеМО. Для понимания процесса составления графа переходов вместо номеров состояний в вершинах графа указаны коды состояний.

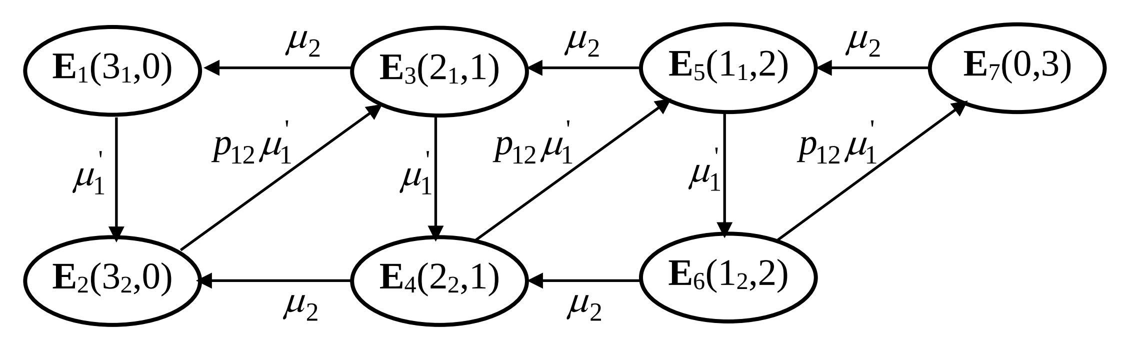


Рис. 5.21. Граф переходов ЗСеМО с эрланговским обслуживанием

Из состояния **E**$_1$=(3_1, 0) переход возможен только в одно состояние **E**$_2$=(3_2, 0) с интенсивностью μ_1' обслуживания на первой фазе, поскольку все заявки в первом узле обязательно проходят две фазы обслуживания.

Из состояния **E**$_2$=(3_2, 0) переход возможен также только в одно состояние **E**$_3$=(2_1, 1). Это соответствует завершению обслуживания на второй фазе заявки в узле 1 (с интенсивностью μ_1') и ее передаче в узел 2 (с вероятностью p_{12}). Отсюда интенсивность перехода марковского процесса в состояние **E**$_3$=(2_1, 1) будет равна произведению $p_{12}\mu_1'$. Заметим, что с вероятностью $(1-p_{12})$ марковский процесс останется в том же состоянии, что соответствует возврату заявки в узел 1.

Из состояния **E**$_3$=(2_1, 1) переход возможен в одно из двух состояний:

- в состояние **E**$_4$=(2_2, 1), что соответствует завершению обслуживания заявки на первой фазе в узле 1 (с интенсивностью μ_1') и переходу к обслуживанию на второй фазе в том же узле 1;

- в состояние **E**$_1$=(3_1, 0), что соответствует завершению обслуживания заявки в узле 2 (с интенсивностью μ_2) и ее передаче в узел 1 (с вероятностью 1).

Следует помнить, что в любой момент времени может произойти только одно событие. Вероятность двух и более событий пренебрежимо мала. Поэтому из состояния **E**$_3$=(2$_1$, 1) не возможен переход в состояние **E**$_3$=(3$_2$, 0), означающий завершение обслуживания заявки на первой фазе в узле 1 и одновременное завершение обслуживания заявки в узле 2.

Аналогичные рассуждения справедливы для состояний **E**$_4$=(2$_2$, 1), **E**$_5$=(1$_1$, 2) и **E**$_6$=(1$_2$, 2).

Из состояния **E**$_7$=(0, 3) переход возможен только в состояние **E**$_5$=(1$_1$, 2), означающий завершение обслуживания заявки в узле 2 и ее передачу на обслуживание в узел 1, причем обслуживание новой заявки всегда начинается на первой фазе. По этой причине переход в состояние **E**$_6$=(1$_2$, 2) не возможен.

6. Система уравнений.

Не составляя матрицу интенсивностей переходов, запишем систему уравнений для определения стационарных вероятностей состояний:

$$\begin{cases} \mu_1' p_1 = \mu_2 p_3 \\ p_{12}\mu_1' p_2 = \mu_1' p_1 + \mu_2 p_4 \\ (\mu_1' + \mu_2) p_3 = p_{12}\mu_1' p_2 + \mu_2 p_5 \\ (p_{12}\mu_1' + \mu_2) p_4 = \mu_1' p_3 + \mu_2 p_6 \\ (\mu_1' + \mu_2) p_5 = p_{12}\mu_1' p_4 + \mu_2 p_7 \\ (p_{12}\mu_1' + \mu_2) p_6 = \mu_1' p_5 \\ \mu_2 p_7 = p_{12}\mu_1' p_6 \\ p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 = 1 \end{cases}$$

7. Расчет характеристик СеМО.

Характеристики ЗСеМО определяются в такой последовательности:

1) загрузка и коэффициенты простоя узлов:

$\rho_1 = p_1 + p_2 + p_3 + p_4 + p_5 + p_6;\quad \rho_2 = p_3 + p_4 + p_5 + p_6 + p_7;$

$\eta_1 = 1 - \rho_1;\quad \eta_2 = 1 - \rho_2;$

2) среднее число параллельно работающих *узлов* сети, определяемое как суммарная *загрузка* всех узлов СеМО:

$R = \rho_1 + \rho_2;$

3) среднее число заявок в очередях и в узлах СеМО:

$l_1 = 2(p_1 + p_2) + p_3 + p_4;\quad l_2 = p_5 + p_6 + 2p_7;$

$m_1 = 3(p_1 + p_2) + 2(p_3 + p_4) + p_5 + p_6;\quad m_2 = p_3 + p_4 + 2(p_5 + p_6) + 3p_7;$

4) суммарное число заявок во всех очередях СеМО:

$L = l_1 + l_2;$

5) производительность замкнутой СеМО:

$$\lambda_0 = \frac{\rho_1}{\alpha_1 b_1} = \frac{\rho_2}{\alpha_2 b_2};$$

где α_1 и α_2 – коэффициенты передач соответственно узла 1 и узла 2;

6) средние времена ожидания и пребывания заявок в узлах СеМО:

$$w_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad w_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

$$u_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad u_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

7) суммарное (полное) время ожидания и время пребывания заявок в СеМО:

$$W = \alpha_1 w_1 + \alpha_2 w_2;$$

$$U = \alpha_1 u_1 + \alpha_2 u_2;$$

8) нагрузка в узлах сети:

$$y_1 = \alpha_1 \lambda_0 b_1; \quad y_2 = \alpha_2 \lambda_0 b_2;$$

9) среднее число параллельно работающих *приборов* во всех узлах сети, определяемое как суммарная *нагрузка* всех узлов СеМО:

$$Y = y_1 + y_2;$$

Суммарное число заявок, циркулирующих в СеМО, рассчитываемое как $M = m_1 + m_2$, должно совпадать с заданным числом заявок в замкнутой сети: $M = 3$.

5.5.4. Замкнутая СеМО с гиперэкспоненциальным обслуживанием

1. Описание СеМО.

1.1. Сеть массового обслуживания (СеМО) – *двухузловая*.

1.2. Количество приборов в узлах: $K_1 = K_2 = 1$.

1.3. Поток заявок *однородный*.

1.4. В СеМО постоянно циркулируют *M*=3 заявки.

2. Предположения и допущения.

2.1. Длительность обслуживания заявок *в узле 1* распределена по *гиперэкспоненциальному закону* со средней длительностью обслуживания заявок $b_1 = 1/\mu_1$ и коэффициентом вариации $\nu_{b_1} = 2$, а в узле 2 – по экспоненциальному закону со средней длительностью обслуживания заявок $b_2 = 1/\mu_2$, где μ_1, μ_2 – интенсивности обслуживания заявок.

2.2. Заявка после обслуживания в узле 1 с вероятностью p_{12} переходит в узел 2 и с вероятностью $p_{10} = 1 - p_{12}$ возвращается в этот же узел 1.

2.3. Дуга, выходящая из узла 1 и входящая обратно в этот же узел, рассматривается как внешняя по отношению к СеМО, и на ней выбирается нулевая точка «0».

В замкнутой СеМО всегда существует стационарный режим.

3. Сведение случайного процесса к марковскому.

Для описания процесса функционирования в замкнутой неэкспоненциальной сети в терминах марковских случайных процессов, как и ранее, будем рассматривать функционирование системы в определенные моменты времени, в которые случайный процесс обладает марковским свойством. Для этого воспользуемся представлением случайной величины, распределенной по гиперэкспоненциальному закону, в виде композиции двух экспоненциально распределенных случайных величин (см. раздел 2, п.2.6.2), каждая из которых появляется с вероятностями q и $(1-q)$ соответственно. В первом узле ЗСеМО такое представление реализуется в виде двух параллельных экспоненциальных фаз, обслуживающих заявки по следующей схеме (рис.5.22):

- заявка с вероятностью $q = 0{,}1$ попадает на обслуживание в первую фазу, длительность обслуживания в которой распределена по экспоненциальному закону со средним значением b_1', после чего покидает узел;
- заявка с вероятностью $(1-q) = 0{,}9$ попадает на обслуживание во вторую фазу, длительность обслуживания в которой распределена по экспоненциальному закону со средним значением b_1'', после чего покидает первый узел.

Значения длительностей обслуживания в этих двух фазах таковы, что выполняется условие: $qb_1' + (1-q)b_1'' = b_1$. Последнее необходимо для того, чтобы средняя длительность обслуживания в узле 1 была равна b_1 .

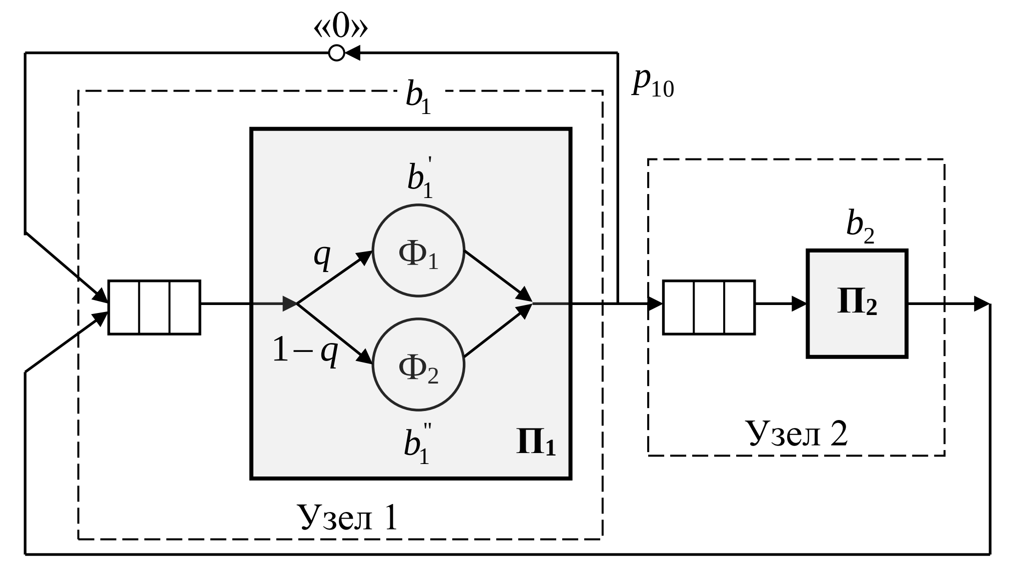


Рис.5.22. ЗСеМО с двухфазным представлением в узле 1 гиперэкспоненциального распределения длительности обслуживания

Моменты завершения обслуживания в каждой из фаз образуют цепь Маркова, так как времена нахождения в них распределены по экспоненциальному закону.

4. Кодирование состояний случайного процесса.

Под состоянием марковского процесса будем понимать распределение заявок по узлам СеМО с учетом того, на какой фазе обслуживания в узле 1 находится заявка.

Для этого закодируем состояния следующим образом: (**M₁, M₂**), где **M₁** = {0, 1_1, 1_2, 2_1, 2_2, 3} – количество заявок, находящихся в узле 1 (индексы отражают нахождение заявки на 1-й или 2-й фазе гиперэкспоненциального распределения), и **M₂** = {0, 1, 2, 3} – количество заявок, находящихся в узле 2, причем суммарное число заявок в обоих узлах должно быть равно 3.

При выбранном способе кодирования система может находиться, как и в предыдущем примере, в следующих состояниях:

E$_1$: (3_1, 0) – все три заявки находятся в узле 1, причем одна заявка находятся на обслуживании в приборе на *первой фазе*, и две заявки ожидают в накопителе;

E$_2$: (3_2, 0) – все три заявки находятся в узле 1, причем одна заявка находятся на обслуживании в приборе на *второй фазе*, и две заявки ожидают в накопителе;

E$_3$: (2_1, 1) – две заявки находятся в узле 1 (одна на обслуживании в приборе на *первой фазе* и одна в накопителе) и одна – на обслуживании в узле 2;

E$_4$: (2_2, 1) – две заявки находятся в узле 1 (одна на обслуживании в приборе на *второй фазе* и одна в накопителе) и одна – на обслуживании в узле 2;

E$_5$: (1_1, 2) – одна заявка находится в узле 1 на обслуживании в приборе на *первой фазе* и две заявки находятся в узле 2, причем одна из них находится на обслуживании в приборе, а вторая заявка ожидает в накопителе;

E$_6$: (1_2, 2) – одна заявка находится в узле 1 на обслуживании в приборе на *второй фазе* и две заявки находятся в узле 2, причем одна из них находится на обслуживании в приборе, а вторая заявка ожидает в накопителе;

E$_7$: (0, 3) – три заявки находятся в узле 2, причем одна заявка – на обслуживании в приборе, а две другие – ожидают в накопителе.

5. Размеченный граф переходов случайного процесса.

На рис.5.23 представлен граф переходов марковского процесса для рассматриваемой неэкспоненциальной СеМО с гиперэкспоненциальным распределением длительности обслуживания заявок в первом узле. Для понимания процесса составления графа переходов вместо номеров состояний в вершинах графа указаны коды состояний, а для того чтобы не загромождать рисунок, используются следующие обозначения для интенсивностей переходов: $g_1 = (1-q)(1-p_{12})\mu_1^{'}$; $g_2 = (1-q)\,p_{12}\mu_1^{'}$; $g_3 = q(1-p_{12})\mu_1^{''}$; $g_4 = qp_{12}\mu_1^{''}$.

Рассмотрим подробно все возможные переходы для каждого состояния $\mathbf{E}_i$ $(i=\overline{1,7})$ марковского случайного процесса.

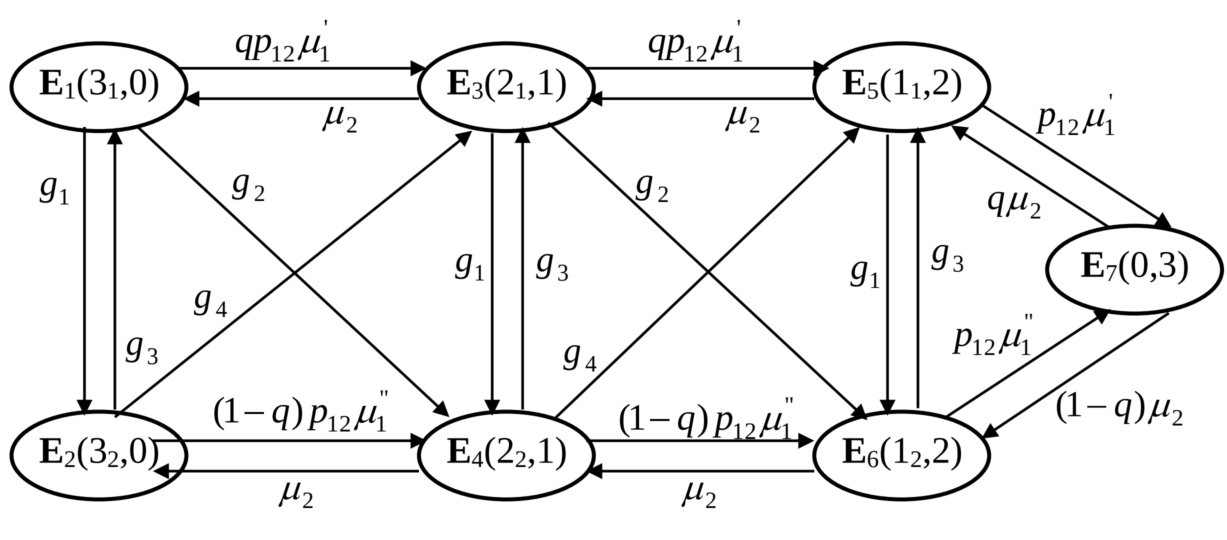


Рис.5.23. Граф переходов ЗСеМО с гиперэкспоненциальным обслуживанием

Состояние $\mathbf{E}_1$. Если случайный процесс находится в состоянии $\mathbf{E}_1$=(3_1, 0), то по завершению обслуживания заявки случайный процесс может перейти в одно из трёх состояний: $\mathbf{E}_2$=(3_2, 0), $\mathbf{E}_3$=(2_1, 1) и $\mathbf{E}_4$=(2_2, 1) или остаться в том же состоянии. Напомним, что если случайный процесс остаётся в том же состоянии, то это никак не отображается на графе переходов.

Случайный процесс перейдёт из состояния $\mathbf{E}_1$=(3_1, 0) в состояние $\mathbf{E}_2$=(3_2, 1) при выполнении следующих условий:

•завершится обслуживание заявки, находящейся на обслуживании в фазе $Ф_1$; интенсивность этого события $\mu_1^{'} = 1/b_1^{'}$;

•заявка, завершившая обслуживание в узле 1, вернётся в этот же узел и встанет в конец очереди; вероятность этого события равна $p_{10} = 1 - p_{12}$;

•в узле 1 очередная заявка, которая поступит на обслуживание из очереди в прибор $П_1$, попадёт на обслуживание в фазу $Ф_2$; вероятность этого события равна $(1-q)$.

Таким образом, интенсивность перехода из состояния $\mathbf{E}_1$=(3_1, 0) в состояние $\mathbf{E}_2$=(3_2, 0) будет равна $g_1 = (1-q)(1-p_{12})\mu_1^{'}$.

Случайный процесс перейдёт из состояния $\mathbf{E}_1$=(3_1, 0) в состояние $\mathbf{E}_3$=(2_1, 1) при выполнении следующих условий:

•завершится обслуживание заявки, находящейся на обслуживании в фазе $Ф_1$; интенсивность этого события $\mu_1^{'} = 1/b_1^{'}$;

•заявка, завершившая обслуживание в узле 1, перейдёт в узел 2; вероятность этого события равна p_{12};

•в узле 1 новая заявка, которая поступит на обслуживание из очереди в прибор $П_1$, попадёт на обслуживание в фазу $Ф_1$; вероятность этого события – q.

Таким образом, интенсивность перехода из состояния $\mathbf{E}_1$=(3_1, 0) в состояние **E**$_3$=(2$_1$, 1) будет равна $qp_{12}\mu_1^{'}$.

Случайный процесс перейдёт из состояния **E**$_1$=(3$_1$, 0) в состояние **E**$_4$=(2$_2$, 1) при выполнении следующих условий:

•завершится обслуживание заявки, находящейся на обслуживании в фазе Ф$_1$; интенсивность этого события $\mu_1^{'} = 1/b_1^{'}$;

•заявка, завершившая обслуживание в узле 1, перейдёт в узел 2; вероятность этого события равна p_{12};

•в узле 1 новая заявка, которая поступит на обслуживание из очереди в прибор П$_1$, попадёт на обслуживание в фазу Ф$_2$; вероятность этого события – $(1-q)$.

Таким образом, интенсивность перехода из состояния **E**$_1$=(3$_1$, 0) в состояние **E**$_4$=(2$_2$, 1) будет равна $g_2 = (1-q)p_{12}\mu_1^{'}$.

Состояние E$_2$. Случайный процесс из состояния **E**$_2$=(3$_2$, 0) по завершению обслуживания заявки также может перейти в одно из трёх состояний: **E**$_1$=(3$_1$, 0), **E**$_3$=(2$_1$, 1) и **E**$_4$=(2$_2$, 1) или остаться в том же состоянии.

Случайный процесс перейдёт из состояния **E**$_2$=(3$_2$, 0) в состояние **E**$_1$=(3$_1$, 1) при выполнении следующих условий:

•с интенсивностью $\mu_1^{''} = 1/b_1^{''}$ завершится обслуживание заявки в фазе Ф$_2$;

•с вероятностью $p_{10} = 1 - p_{12}$ заявка, завершившая обслуживание в узле 1, вернётся в этот же узел и встанет в конец очереди;

•с вероятностью q в узле 1 очередная заявка, которая поступит из очереди в прибор П$_1$, попадёт на обслуживание в фазу Ф$_1$.

Таким образом, интенсивность перехода из состояния **E**$_1$=(3$_1$, 0) в состояние **E**$_2$=(3$_2$, 0) будет равна $g_3 = q(1-p_{12})\mu_1^{''}$.

Случайный процесс перейдёт из состояния **E**$_2$=(3$_2$, 0) в состояние **E**$_3$=(2$_1$, 1) при выполнении следующих условий:

•с интенсивностью $\mu_1^{''} = 1/b_1^{''}$ завершится обслуживание заявки в фазе Ф$_2$;

•с вероятностью p_{12} заявка, завершившая обслуживание в узле 1, перейдёт в узел 2;

•с вероятностью q в узле 1 очередная заявка, которая поступит из очереди в прибор П$_1$, попадёт на обслуживание в фазу Ф$_1$.

Таким образом, интенсивность перехода из **E**$_2$=(3$_2$, 0) в **E**$_3$=(2$_1$, 1) будет равна $g_4 = qp_{12}\mu_1^{''}$.

Случайный процесс перейдёт из состояния **E**$_2$=(3$_2$, 0) в состояние **E**$_4$=(2$_2$, 1) при выполнении следующих условий:

•с интенсивностью $\mu_1^{''} = 1/b_1^{''}$ завершится обслуживание заявки в фазе Ф$_2$;

•с вероятностью p_{12} заявка, завершившая обслуживание в узле 1, перейдёт в узел 2;

•с вероятностью $(1-q)$ в узле 1 очередная заявка, которая поступит из очереди в прибор $П_1$, попадёт на обслуживание в фазу $Ф_2$.

Таким образом, интенсивность перехода из **E**$_2$=(3_2, 0) в **E**$_4$=(2_2, 1) будет равна $(1-q)p_{12}\mu_1^{''}$.

Состояния $\mathbf{E}_3$ и $\mathbf{E}_4$. Если случайный процесс находится в состоянии **E**$_3$=(2_1, 1) или **E**$_4$=(2_2, 1), то кроме аналогичных переходов, связанных с завершением обслуживания заявки в узле 1, имеется ещё один переход в состояния **E**$_1$=(3_1, 0) и **E**$_2$=(3_2, 0) соответственно, связанный с завершением обслуживания заявки в узле 2. Интенсивность перехода из **E**$_3$=(2_1, 1) в **E**$_1$=(3_1, 0) и из **E**$_4$=(2_2, 1) в **E**$_2$=(3_2, 0) равна интенсивности обслуживания μ_2 в узле 2. Отметим, что переходы из **E**$_3$=(2_1, 1) в **E**$_2$=(3_2, 0) и из **E**$_4$=(2_2, 1) в **E**$_1$=(3_1, 0) отсутствуют, так как заявка, находящаяся на обслуживании в первом узле, остаётся в той же фазе обслуживания, которая была в момент завершения обслуживания заявки в узле 2. Это является следствием того, что в случайных процессах с непрерывным временем вероятность одновременного появления двух событий (завершение обслуживания в узле 1 и в узле 2) равна нулю.

Состояния $\mathbf{E}_5$ и $\mathbf{E}_6$. Переходы из состояний **E**$_5$=(1_1, 2) и **E**$_6$=(1_2, 2) аналогичны переходам из **E**$_3$=(2_1, 1) и **E**$_4$=(2_2, 1) за исключением переходов в состояние **E**$_7$=(0, 3). Интенсивности переходов из **E**$_5$=(1_1, 2) и **E**$_6$=(1_2, 2) в **E**$_7$=(0, 3) определяются как произведение интенсивности обслуживания в соответствующей фазе узла 1 на вероятность того, что заявка, завершившая обслуживание в узле 1, перейдёт в узел 2: $p_{12}\mu_1^{'}$ и $p_{12}\mu_1^{''}$.

Состояние $\mathbf{E}_7$. Переходы из состояния **E**$_7$=(0, 3) связаны с завершением обслуживания с интенсивностью μ_2 заявки в узле 2, которая переходит в узел 1 и с вероятностью q попадает на обслуживание в фазу $Ф_1$ или с вероятностью $(1-q)$ – в фазу $Ф_2$. Соответственно интенсивности переходов будут равны $q\mu_2$ и $(1-q)\mu_2$

6. Расчет характеристик СеМО.

Не составляя матрицу интенсивностей переходов и не выписывая систему линейных алгебраических уравнений для определения стационарных вероятностей состояний, приведём математические зависимости для расчёта характеристик функционирования ЗСеМО:

1) загрузка и коэффициенты простоя узлов:

$\rho_1 = p_1 + p_2 + p_3 + p_4 + p_5 + p_6$; $\rho_2 = p_3 + p_4 + p_5 + p_6 + p_7$;

$\eta_1 = 1-\rho_1$; $\eta_2 = 1-\rho_2$;

2) среднее число параллельно работающих *узлов* сети, определяемое как суммарная *загрузка* всех узлов СеМО:

$R = \rho_1 + \rho_2$;

3) среднее число заявок в очередях и в узлах СеМО:

$l_1 = 2(p_1 + p_2) + p_3 + p_4; \quad l_2 = p_5 + p_6 + 2p_7;$

$m_1 = 3(p_1 + p_2) + 2(p_3 + p_4) + p_5 + p_6;$

$m_2 = p_3 + p_4 + 2(p_5 + p_6) + 3p_7;$

4) суммарное число заявок во всех очередях СеМО:

$L = l_1 + l_2;$

5) производительность замкнутой СеМО:

$$\lambda_0 = \frac{\rho_1}{\alpha_1 b_1} = \frac{\rho_2}{\alpha_2 b_2};$$

где α_1 и α_2 - коэффициенты передачи соответственно узла 1 и узла 2;

6) средние времена ожидания и пребывания заявок в узлах СеМО:

$$w_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad w_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

$$u_1 = \frac{l_1}{\alpha_1 \lambda_0}; \quad u_2 = \frac{l_2}{\alpha_2 \lambda_0};$$

7) суммарное (полное) время ожидания и время пребывания заявок в СеМО:

$W = \alpha_1 w_1 + \alpha_2 w_2;$

$U = \alpha_1 u_1 + \alpha_2 u_2;$

8) нагрузка в узлах сети:

$y_1 = \alpha_1 \lambda_0 b_1; \quad y_2 = \alpha_2 \lambda_0 b_2;$

9) среднее число параллельно работающих *приборов* во всех узлах сети, определяемое как суммарная *нагрузка* всех узлов СеМО:

$Y = y_1 + y_2$.

Суммарное число заявок, циркулирующих в СеМО, рассчитываемое как $M = m_1 + m_2$, должно совпадать с заданным числом заявок в замкнутой сети: $M = 3$.

Задание на самостоятельную работу:

1. По графу переходов рис.5.23 построить матрицу интенсивностей переходов и составить систему линейных алгебраических уравнений для расчёта стационарных вероятностей состояний.

2. Выполнить детальный анализ свойств исследуемой системы.

5.6. Резюме

1. Марковские случайные процессы используются в качестве математических моделей систем со стохастическим характером функционирования. Марковская модель представляется в виде систем дифференциальных и алгебраических уравнений, для решения которых обычно применяются численные методы. Поэтому марковские случайные процессы можно отнести к *численным методам моделирования*.

2. Случайный процесс полностью описывается перечнем *состояний*, которые задаются значениями некоторых переменных, и *переходами* между состояниями.

3. Для с*лучайного процесса с дискретными состояниями* характерен скачкообразный переход из состояния в состояние, которые могут быть пронумерованы. При этом число возможных состояний может быть *конечным* или *бесконечным*. Для *случайного процесса с непрерывными состояниями* характерен плавный переход из состояния в состояние.

4. Случайные процессы с дискретными состояниями делятся на процессы с *дискретным временем*, когда переходы из состояния в состояние возможны в строго *определенные заранее фиксированные моменты времени*, которые можно пронумеровать, и с *непрерывным временем*, когда интервал времени между соседними переходами является *случайным*, и переход возможен в любой заранее не известный момент времени.

5. Случайные процессы с дискретными состояниями изображаются в виде *графа переходов (состояний)*. В *размеченном* графе переходов на дугах графа указываются условия перехода в виде *вероятностей переходов* или *интенсивностей переходов*.

Состояния случайного процесса могут быть *невозвратными* и *поглощающими*.

6. Случайный процесс называется *марковским*, если вероятность любого состояния в будущем зависит только от его состояния в настоящем и не зависит от того, когда и каким образом процесс оказался в этом состоянии. Для того чтобы случайный процесс с непрерывным временем был *марковским*, необходимо, чтобы интервалы времени между соседними переходами из состояния в состояние были распределены *по экспоненциальному закону*, который обладает замечательным *свойством*: если время нахождения случайного процесса в некотором состоянии распределено по экспоненциальному закону, то *интервал от любого случайного момента времени до момента перехода* в другое состояние имеет *то же экспоненциальное распределение с тем же параметром*. Эта особенность является следствием *отсутствия последействия*, присущего процессам с экспоненциальным распределением времени нахождения в том или ином состоянии.

7. Для описания марковского случайного процесса используется следующая совокупность параметров:

- *перечень состояний* $\mathbf{E}_1, ..., \mathbf{E}_n$;

•*матрица переходов*, в виде *матрицы вероятностей переходов* **Q** для процессов *с дискретным временем* или *матрицы интенсивностей переходов* **G** для процессов *с непрерывным временем*;

•*начальные вероятности* $p_1(0),\ldots, p_n(0)$.

8. Для описания переходов между состояниями случайного процесса *с дискретным временем* используется квадратная *матрица вероятностей переходов* $\mathbf{Q}=[q_{ij} \mid i,j=\overline{1,n}]$, элементы которой удовлетворяют условиям:

$$0 \le q_{ij} \le 1; \quad \sum_{j=1}^{n} q_{ij} = 1 \qquad (i,j=\overline{1,n}).$$

Для описания переходов между состояниями случайного процесса *с непрерывным временем* используется квадратная *матрица интенсивностей переходов* $\mathbf{G}=[g_{ij} \mid i,j=\overline{1,n}]$, в которой *интенсивность перехода* g_{ij} определяется как предел отношения вероятности перехода $P_{ij}(\Delta\tau)$ из состояния $\mathbf{E}_i$ в состояние $\mathbf{E}_j$ за промежуток времени $\Delta\tau$ к длине этого промежутка:

$$g_{ij} = \lim_{\Delta\tau\to 0} \frac{P_{ij}(\Delta\tau)}{\Delta\tau} \qquad (i,j=\overline{1,n};\ i\ne j),$$

а диагональные элементы определяются из условия:

$$\sum_{j=1}^{n} g_{ij} = 0 \quad (i=\overline{1,n}).$$

9. Изучение случайных процессов заключается в определении вероятностей состояний $p_1(t),\ldots,p_n(t)$, которые могут быть представлены *стохастическим* вектором:

$$P(t)=\{p_1(t),\ldots,p_n(t)\},$$

причем

$$0 \le p_i(t) \le 1; \quad \sum_{i=1}^{n} p_i(t) = 1.$$

Вектор состояний $P(t)=\{p_1(t),\ldots,p_n(t)\}$ является *основной* характеристикой марковского случайного процесса.

10. Случайный процесс обладает *эргодическим свойством*, если по истечении достаточно большого промежутка времени вероятности состояний стремятся к предельным (стационарным) значениям $p_1,\ldots,p_n$, не зависящим от начальных вероятностей $p_1(0),\ldots,p_n(0)$ и от самого промежутка времени. В этом случае система, в которой протекает случайный процесс, работает в *установившемся* или *стационарном режиме*. В противном случае система работает в нестационарном режиме.

Случайный процесс *с дискретным временем* обладает *эргодическим свойством*, если матрица вероятностей переходов *не является периодической* или *разложимой*. Случайный процесс с *непрерывным временем* и *конечным* числом состояний всегда обладает эргодическим свойством.

11. Для марковского процесса с дискретным временем, обладающего эргодическим свойством, стационарные вероятности состояний определяются из системы линейных алгебраических уравнений:

$$p_j = \sum_{i=1}^{n} p_i q_{ij} \quad (j = \overline{1,n}),$$

которая совместно с нормировочным условием $\sum_{i=1}^{n} p_i = 1$ образует систему, обладающую единственным решением.

Аналогично, для марковского процесса с непрерывным временем, обладающего эргодическим свойством, стационарные вероятности состояний определяются из системы линейных алгебраических уравнений:

$$\sum_{i=1}^{n} p_i g_{ij} = 0 \quad (j = \overline{1,n}),$$

которая совместно с нормировочным условием образует систему, обладающую единственным решением.

5.7. Практикум: обсуждение и решение задач

Вопрос 1. Существуют ли реальные системы, в которых протекающие в них случайные процессы являются марковскими?

Обсуждение. Марковский процесс является такой же идеализированной моделью реальных систем, как и простейший поток, представляющий собой идеализированную модель случайного потока заявок. Эта идеализация заключается в том, что с вероятностью, отличной от нуля, марковский процесс может находиться в любом из состояний бесконечно долго. Это обусловлено тем, что плотность экспоненциального распределения ограничена слева и не ограничена справа. Очевидно, что в реальных системах это невозможно. В то же время, как показывают многочисленные исследования, такая идеализация часто оказывается оправданной, поскольку при определённых условиях позволяет получить для многих реальных систем вполне приемлемые результаты, погрешность которых лежит в допустимых для практики пределах в 10-20%. Кроме того, в некоторых случаях предположение о марковском характере протекающих в исследуемой системе процессов позволяет получить верхние оценки характеристик функционирования системы.

Вопрос 2. Когда случайный процесс с непрерывным временем не обладает эргодическим свойством?

Обсуждение. Случайный процесс с непрерывным временем не обладает эргодическим свойством, если среди его состояний имеются невозвратные или поглощающие состояния. В первом случае это означает, что по истечении некоторого (иногда достаточно большого) времени случайный процесс никогда не сможет попасть в невозвратные состояния, а во втором случае – процесс окажется в одном из поглощающих состояний, из которого он никогда не сможет выйти.

Вопрос 3. Обладает ли эргодическим свойством случайный процесс с непрерывным временем, имеющий бесконечное число состояний?

Обсуждение. Случайный процесс с непрерывным временем и бесконечным числом состояний может обладать или не обладать эргодическим свойством. Применительно к случайным процессам, протекающим в системах массового обслуживания, наличие эргодического свойства определяется наличием установившегося режима в моделируемой системе, а точнее отсутствием перегрузок в системе с накопителями неограниченной ёмкости. Если же система перегружена, что со временем приводит к бесконечному увеличению длины очереди заявок в системе, то можно утверждать, что соответствующий случайный процесс не будет обладать эргодическим свойством.

Задача 1. Определить, обладает ли эргодическим свойством случайный процесс с дискретным временем с заданной матрицей вероятностей переходов P, сопроводив ответ необходимыми пояснениями.

$$P = \begin{array}{c|cccc|} & E_1 & E_2 & E_3 & E_4 \\ E_1 & 0 & 0.2 & 0 & 0.8 \\ E_2 & 0.1 & 0.2 & 0.3 & 0.4 \\ E_3 & 0 & 0.5 & 0 & 0.5 \\ E_4 & 0.2 & 0.4 & 0 & 0.4 \end{array}$$

Решение. Случайный процесс обладает эргодическим свойством, если матрица вероятностей переходов не является разложимой или периодической. Переставляя столбцы и строки матрицы, проверим, является ли заданная матрица P разложимой или периодической.

Рассмотрим два варианта перестановок столбцов и строк:

$$P_1 = \begin{array}{c|cccc|} & E_1 & E_3 & E_2 & E_4 \\ E_1 & 0 & 0 & 0.2 & 0.8 \\ E_3 & 0 & 0 & 0.5 & 0.5 \\ E_2 & 0.1 & 0.3 & 0.2 & 0.4 \\ E_4 & 0.2 & 0 & 0.4 & 0.4 \end{array} \qquad P_2 = \begin{array}{c|cccc|} & E_2 & E_4 & E_1 & E_3 \\ E_2 & 0.3 & 0.4 & 0.1 & 0.3 \\ E_4 & 0.4 & 0.4 & 0.2 & 0 \\ E_1 & 0.2 & 0.8 & 0 & 0 \\ E_3 & 0.5 & 0.5 & 0 & 0 \end{array}$$

Полученные матрицы P_1 и P_2 с нулевыми подматрицами в верхнем левом углу в P_1 и в нижнем правом углу в P_2 не являются разложимыми или периодическими, следовательно, случайный процесс обладает эргодическим свойством

Задача 2. Известны вероятности состояний трехузловой замкнутой экспоненциальной СеМО: P(0,0,2)=0,3; P(0,1,1)=0,4; P(0,2,0)=0,1; P(1,0,1)=0,05; P(1,1,0)=0,05; Р(2,0,0)=0,1. Длительности обслуживания заявок во всех одноканальных узлах одинаковы. Определить значения коэффициентов передач второго и третьего узлов сети, если известно, что коэффициент передачи первого узла равен 4.

Дано: ЗСеМО: $n = 3$; $K_1 = K_2 = K_3 = 1$;

$b_1 = b_2 = b_3 = b$; $\alpha_1 = 4$;

$P(0,0,2) = 0,3$; $P(0,1,1) = 0,4$; $P(0,2,0) = 0.1$;

$P(1,0,1) = 0,05$; $P(1,1,0) = 0,05$; $P(2,0,0) = 0.1$.

Определить: $\alpha_2 = ?$ и $\alpha_3 = ?$

Решение.

1) По заданным значениям стационарных вероятностей состояний с учётом того, что все узлы одноканальные, рассчитаем загрузки каждого узла замкнутой СеМО как сумму вероятностей состояний, в которых соответствующий узел занят обслуживанием заявок:

$$\rho_1 = P(1,0,1) + P(1,1,0) + P(2,0,0) = 0.05 + 0.05 + 0.1 = 0.2;$$

$$\rho_2 = P(0,1,1) + P(0,2,0) + P(1,1,0) = 0.4 + 0.1 + 0.05 = 0.55;$$

$$\rho_3 = P(0,0,2) + P(0,1,1) + P(1,0,1) = 0.3 + 0.4 + 0.05 = 0.75.$$

2) Загрузка узлов СеМО (см.п.3.4.3) определяется по формуле:

$$\rho_j = \frac{\alpha_j \lambda_0 b_j}{K_j} \quad (j = \overline{1,3})$$

или с учётом того, что $K_1 = K_2 = K_3 = 1$ и $b_1 = b_2 = b_3 = b$, получим:

$$\rho_j = \alpha_j \lambda_0 b \quad (j = \overline{1,3}),$$

где λ_0 - интенсивность потока заявок, проходящих через нулевой узел ЗСеМО, значение которой не известно.

Зная загрузку $\rho_1 = 0.2$ и коэффициент передачи $\alpha_1 = 4$ узла 1, найдём:

$$\lambda_0 b = \rho_1 / \alpha_1 = 0.2 / 4 = 0.05.$$

3) Теперь с использованием того же выражения для расчёта загрузок узлов 2 и 3 можно определить значения соответствующих коэффициентов передач:

$$\alpha_2 = \frac{\rho_2}{\lambda_0 b} = \frac{0,55}{0,05} = 11; \quad \alpha_3 = \frac{\rho_3}{\lambda_0 b} = \frac{0,75}{0,05} = 25.$$

Задача 3. На автозаправочную станцию (АЗС) с одной колонкой прибывают автомобили со средним интервалом между моментами прибытия *X* минут. Водитель каждого автомобиля сначала заправляет бензином автомобиль в течение случайного времени, распределённого по экспоненциальному закону, со средним значением Y минут, а затем идёт к оператору АЗС и оплачивает бензин, затрачивая на это в среднем ещё Y минут. После этого автомобиль покидает заправку, и к колонке подъезжает следующий ожидающий заправки автомобиль. Ожидающие автомобили образуют очередь перед АЗС.

1) Сформулировать предположения и допущения, при которых процесс функционирования бензозаправочной станции можно рассматри-

вать как марковский.

2) Нарисовать и подробно описать модель в терминах теории массового обслуживания.

3) Выполнить кодирование и нарисовать размеченный граф переходов марковского процесса.

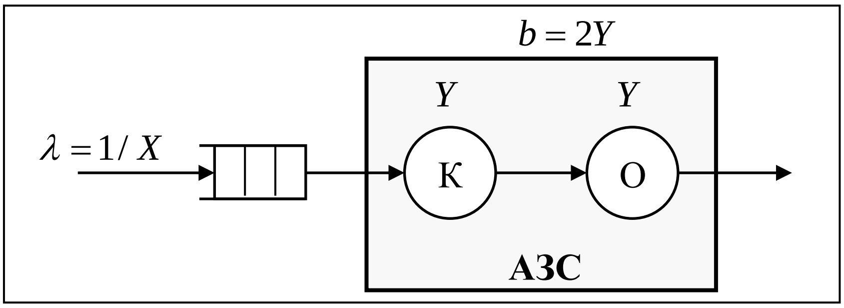
4) Сформулировать требования, при которых марковский процесс обладает эргодическим свойством.

Решение.

1) *Предположения и допущения*, при которых процесс функционирования бензозаправочной станции можно рассматривать как марковский:

- прибывающие на бензозаправочную станцию автомобили образуют *простейший поток*;
- время, затрачиваемое на заправку, и время, затрачиваемое на оплату за бензин, представляют собой случайные величины, распределённые по *экспоненциальному закону*;
- интервал времени от момента отъезда от бензоколонки заправленного автомобиля до момента подъезда к бензоколонке следующего ожидающего автомобиля предполагается много меньшим по сравнению со временем заправки и принимается равным *нулю*;
- в очереди ожидающих заправки автомобилей может находиться любое их количество, то есть имеем накопитель *неограниченной ёмкости*.

2) *Модель* в терминах теории массового обслуживания:



Модель АЗС представляет собой *одноканальную* СМО с накопителем *неограниченной* ёмкости, в которую поступает *простейший* поток заявок (автомобилей) с интенсивностью $\lambda = 1/X$. Обслуживание в приборе складывается из *двух экспоненциальных фаз*: на первой фазе (К) выполняется заправка на колонке автомобиля бензином, а на второй (О) – оплата за бензин. Интенсивность обслуживания на каждой фазе равна $\mu = 1/Y$ заявок в минуту, следовательно, интенсивность обслуживания в приборе (АЗС) составляет $1/(2Y) = \mu/2$. Предположение об экспоненциальном характере обслуживания на каждой фазе обусловливает распределение длительности обслуживания в приборе по *закону Эрланга 2-го порядка*.

3) *Кодирование* и *размеченный граф* переходов марковского процесса.

В качестве параметра, описывающего состояние марковского процесса, будем рассматривать количество заявок k, находящихся в СМО (на обслуживании в приборе и в накопителе), при этом следует различать, на какой экспоненциальной фазе обслуживания в приборе находится заявка. Поскольку в системе в произвольный момент времени может находиться любое сколь угодно большое число заявок, то количество состояний марковского процесса равно бесконечности:

E$_0$: $k = 0$ – в системе нет ни одной заявки;

E$_1$: $k = 1_1$ – в системе находится 1 заявка на обслуживании в фазе 1;

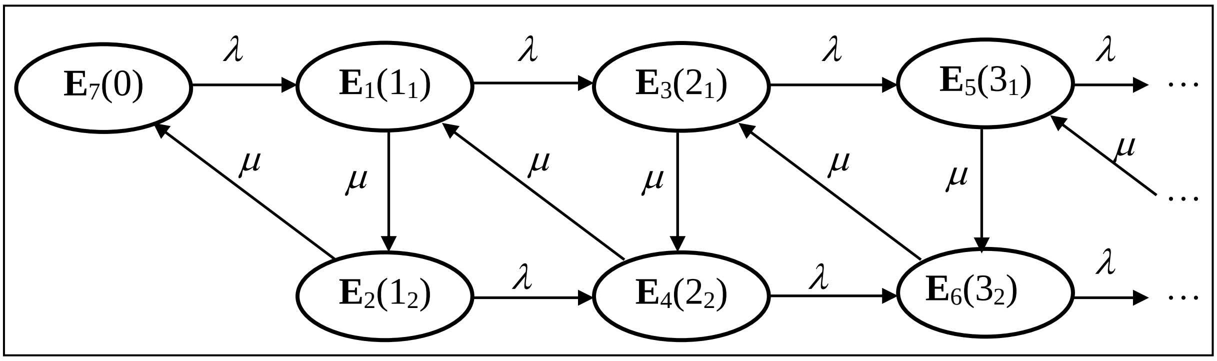
E$_2$: $k = 1_2$ – в системе находится 1 заявка на обслуживании в фазе 2;

E$_3$ $k = 2_1$ – в системе находятся 2 заявки (одна – на обслуживании в фазе 1 и вторая ожидает в накопителе);

E$_4$ $k = 2_2$ – в системе находятся 2 заявки (одна – на обслуживании в фазе 2 и вторая ожидает в накопителе);

…

Размеченный граф переходов имеет следующий вид:



4) *Требования*, при которых марковский процесс обладает эргодическим свойством.

Марковский процесс с непрерывным временем и бесконечным количеством состояний обладает эргодическим свойством, если в моделируемой системе нет перегрузок. Для этого необходимо, чтобы загрузка системы не превышала единицы:

$$\rho = \lambda b = \frac{2Y}{X} < 1.$$

Отсюда вытекает очевидное требование следующего вида: $X > 2Y$, то есть средний интервал между прибывающими на АЗС автомобилями должен быть больше, чем среднее время их обслуживания, затрачиваемое на заправку и оплату.

Если это условие не выполняется, можно ограничить ёмкость накопителя, построив перед АЗС площадку с ограниченным числом мест для ожидающих автомобилей, полагая, что при отсутствии на этой площадке свободных мест автомобили отправятся на другую АЗС.

5.8. Самоконтроль: перечень вопросов и задач

1. Понятие случайного процесса.
2. Что понимается под состоянием случайного процесса?
3. Классификация случайных процессов.
4. В чём отличие дискретного случайного процесса от непрерывного?
5. Привести примеры систем, в которых процессы непрерывными.
6. Привести примеры систем, в которых процессы дискретными.
7. В чём отличие дискретного случайного процесса с непрерывным временем от процесса с дискретным временем?
8. Понятие марковского случайного процесса.
9. Как называется процесс, в котором переход из одного состояния в другое зависит только от состояния, в котором находится процесс?
10. При каком условии случайный процесс с непрерывным временем является марковским?
11. По какому закону должны быть распределены интервалы времени между соседними переходами, чтобы дискретный случайный процесс был марковским? Ответ обосновать.
12. Дать определение интенсивности перехода для марковского случайного процесса с непрерывным временем.
13. Из какого условия определяются диагональные элементы матрицы интенсивностей переходов?
14. Чему равны диагональные элементы матрицы интенсивностей переходов?
15. Понятие эргодического свойства случайного процесса.
16. В чем различие между случайными процессами, обладающими и не обладающими эргодическим свойством?
17. Что означает понятие "стационарная вероятность состояния случайного процесса"?
18. Перечислить условия, при которых марковский процесс с дискретным временем обладает эргодическим свойством.
19. Объяснить на примере, почему марковский процесс с разложимой и периодической матрицей вероятностей переходов не обладает эргодическим свойством?
20. Определить, обладает ли эргодическим свойством случайный процесс с дискретным временем с заданной матрицей вероятностей переходов P, сопроводив ответ необходимыми пояснениями.

$$P = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \\ \end{array}$$

$$P = \begin{array}{c|cccc|} & E_1 & E_2 & E_3 & E_4 \\ E_1 & 0.2 & 0 & 0.8 & 0 \\ E_2 & 0 & 0.2 & 0.3 & 0.4 \\ E_3 & 0.4 & 0 & 0.6 & 0 \\ E_4 & 0.2 & 0.3 & 0 & 0.5 \end{array}$$

21. Известны вероятности состояний двухузловой замкнутой СеМО: P(0,4)=0,4; P(1,3)=0,1; P(2,2)=0,2; P(3,1)=0,2; P(4,0)=0,1, где состояние (M_1,M_2) задает число заявок в одноканальном узле 1 и трехканальном узле

2 соответственно. Определить среднее число заявок в СеМО, находящихся в состоянии ожидания.

22. Известны вероятности состояний трехузловой замкнутой СеМО: P(0,0,2)=0,2; P(0,1,1)=0,1; P(0,2,0)=0,15; P(1,0,1)=0,35; P(1,1,0)=0,15; P(2,0,0)=0,05, где состояние (M_1,M_2,M_3) задает число заявок в узле 1, 2, 3 соответственно. Определить среднее число параллельно работающих узлов ЗСеМО.

23. Известны вероятности состояний трехузловой замкнутой СеМО: P(0,0,2)=0,1; P(0,1,1)=0,3; P(0,2,0)=0,4; P(1,0,1)=0,05; P(1,1,0)=0,05; P(2,0,0)=0,1. Длительности обслуживания заявок во всех одноканальных узлах одинаковы. Определить значения коэффициентов передач второго и третьего узлов сети, если известно, что коэффициент передачи первого узла равен 2.

24. Известны вероятности состояний трехузловой ЗСеМО: P(2,0,0)=0,05; P(1,1,0)=0,25; P(0,2,0)=0,1; P(1,0,1)=0,1; P(0,1,1)=0,3; P(0,0,2)=0,2. Определить производительность ЗСеМО, если известно, что коэффициент передачи третьего узла (двухканального) равен 2, а средняя длительность обслуживания заявок в этом узле равна 0,1 с.

25. Система содержит два обслуживающих прибора и накопитель единичной емкости (для одной заявки). В систему поступает простейший поток заявок с интенсивностью λ. Заявки с равной вероятностью попадают в один из них, если оба прибора свободны, и занимают свободный прибор, когда другой прибор занят обслуживанием. Когда оба прибора заняты, заявка заносится в накопитель, если он свободен, или теряется, если накопитель занят. Длительность обслуживания заявок в обоих приборах распределена по гиперэкспоненциальному закону, причем первый прибор работает с двое большей скоростью. Нарисовать модель системы и размеченный граф переходов марковского процесса с необходимыми для понимания комментариями. Составить систему уравнений для стационарных вероятностей .

26. Система содержит два обслуживающих прибора и накопитель единичной емкости (для одной заявки). В систему поступают заявки с интенсивностью λ. Если оба прибора свободны, то поступившая заявка всегда попадает в первый прибор, и занимают свободный прибор, когда другой прибор занят обслуживанием. Когда оба прибора заняты, заявка заносится в накопитель, если он свободен, или теряется, если накопитель занят. Первый прибор работает с вдвое меньшей скоростью. 1) Сформулировать условия (предположения и допущения), при которых случайный процесс, протекающий в системе, будет марковским. 2) Нарисовать модель системы. 3) Выполнить кодирование марковского процесса. 4) Нарисовать размеченный граф переходов марковского процесса. 5) Выписать систему уравнений для определения стационарных вероятностей состояний. 6) Сформулировать условия, при которых марковский процесс обладает эргодическим свойством.

27. На автозаправочной станции (АЗС) имеется две колонки: одна для заправки легковых автомобилей бензином и другая для заправки грузовых автомобилей дизельным топливом. На станцию прибывают автомобили со средним интервалом между моментами прибытия T_0 минут, причём легковые автомобили прибывают в 4 раза чаще, чем грузовые. Время заправки легковых автомобилей в среднем составляет X минут, а грузовых – в два раза больше. Перед АЗС имеется площадка для ожидания прибывающих автомобилей, на которой могут разместиться один грузовой или два легковых автомобиля. Если площадка занята, то автомобили покидают АЗС не заправившись. 1) Сформулировать предположения и допущения, при которых процесс функционирования бензозаправочной станции можно рассматривать как марковский. 2) Нарисовать и подробно описать модель в терминах теории массового обслуживания. 3) Выполнить кодирование и нарисовать размеченный граф переходов марковского процесса. 4) Сформулировать требования, при которых марковский процесс будет обладать эргодическим свойством.

28. В мужской парикмахерской работает один мастер. Средний интервал между моментами прихода клиентов составляет X минут. Каждый клиент просит сначала побрить, а затем постричь. Мастер тратит на каждую из этих операций случайное время со средним значением Y минут. В парикмахерской имеется одно кресло для ожидания. Если кресло занято, то очередной пришедший клиент уходит из парикмахерской не обслуженным. 1) Сформулировать предположения и допущения, при которых процесс функционирования парикмахерской можно рассматривать как марковский. 2) Нарисовать и подробно описать модель в терминах теории массового обслуживания. 3) Выполнить кодирование марковского процесса. 4) Нарисовать размеченный граф переходов марковского процесса. 5) Выписать систему уравнений для определения вероятностей состояний. 6) Сформулировать требования, при которых марковский процесс обладает эргодическим свойством.

29. В парикмахерскую, в которой работают мастер и ученик, приходят клиенты в среднем с интервалом t_1 минут. Пришедший клиент направляется к мастеру, если он свободен, и к ученику, в противном случае. Когда мастер и ученик заняты, клиент располагается в зале на имеющемся там единственном стуле для ожидания, если он свободен. Если стул занят, то пришедший клиент покидает парикмахерскую. Мастер работает вдвое быстрей, чем ученик. 1) Сформулировать условия, при которых процесс функционирования парикмахерской можно представить в виде марковского процесса. 2) Нарисовать детальную модель системы с подробным ее описанием. 3) Выполнить кодирование состояний и нарисовать размеченный граф переходов марковского процесса. 4) Составить систему уравнений для стационарных вероятностей.