Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
ПЕТРОЗАВОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Е. К. Белый

Введение в теорию массового обслуживания

Учебное пособие для студентов, обучающихся по направлению «Информационные системы и технологии»

Петрозаводск
Издательство ПетрГУ
2014

УДК 519.21
ББК 22.172
Д439

Печатается по решению редакционно-издательского совета
Петрозаводского государственного университета

Издается в рамках реализации комплекса мероприятий
Программы стратегического развития ПетрГУ на 2012—2016 гг.

Рецензенты:
д-р физ.-мат. наук, профессор каф. геометрии и топологии ПетрГУ
С. С. Платонов;
д-р эконом. наук, зав. отделом моделирования и прогнозирования
регионального развития института экономики КарНЦ РАН
П. В. Дружинин

Белый, Евгений Константинович.
Б439 Введение в теорию массового обслуживания : учебное пособие для студентов, обучающихся по направлению «Информационные системы и технологии» / Е. К. Белый. – Петрозаводск : Издательство ПетрГУ, 2014. – 76 с.

ISBN 978-5-8021-2203-7

Учебное пособие предназначено для студентов, обучающихся по направлению «Информационные системы и технологии», а также для всех интересующихся теорией массового обслуживания и владеющих математическим аппаратом в рамках вузовского курса высшей математики и теории вероятностей.

УДК 519.21
ББК 22.172

ISBN 978-5-8021-2203-7

.

Содержание

Введение

То и дело раздаются голоса, утверждающие, будто главная задача обучения математике в школе и вузе – это научить людей логически мыслить. Отсюда чрезмерная формализация математических дисциплин, изложение их в отрыве от задач практики. Слов нет, привычка к логическому мышлению – хорошее дело, но у математики есть и другие задачи: активного вмешательства в практику, разумной организации производственных и иных процессов. Жизнь непрерывно требует от математика ответа на вопрос, как поступить в том или другом случае, при тех или других сложившихся обстоятельствах. И дело его чести – не уходить от этих требований в пучину абстракций, а по мере сил удовлетворять их.

Е. С. Вентцель

Теория массового обслуживания родилась в датском королевстве в начале XX века под именем «Теория очередей». Первые идеи теории были высказаны директором Копенгагенской телефонной компании Фредериком Йохансоном в 1907 году в статье «Время ожидания и число вызовов». Затем идеи были математически развиты и оформлены инженером той же компании Агнером Эрлангом. Опубликованную им в 1909 году статью «Теория вероятностей и телефонные переговоры» принято считать краеугольным камнем в фундаменте теории. В 30-х годах теорией очередей серьезно занялся Александр Яковлевич Хинчин в связи с автоматизацией московской городской телефонной сети. В научной литературе прижился введенный тогда Хинчиным термин «теория массового обслуживания» (ТМО), а предмет исследований вскоре стали называть системами массового обслуживания (СМО). ТМО опиралась на фундаментальные работы в области теории случайных процессов Андрея Андреевича Маркова, Андрея Николаевича Колмогорова и ряда других математиков.

Хотя все началось с телефона, вскоре на ТМО обратили внимание представители ряда наук, далеких от проблем связи. Оказалось, что в модели СМО вписываются многие реальные явления: от железнодорожных касс до противовоздушной обороны и процессов, протекающих в компьютере. С системами массового обслуживания мы сталкиваемся буквально на каждом шагу. СМО – это торговля, общественное питание, транспорт, бани, банки, страховые компании, налоговая инспекция, парикмахерские, индустрия развлечений, ремонтные мастерские, здравоохранение, образование и т. д.

Что же представляет собой СМО? Это система, реализующая многократное выполнение однотипных задач. Для любой такой системы характерен, по крайней мере, один входящий поток – поток заявок на обслуживание, множество каналов обслуживания и как минимум один выходящий поток – поток обслуженных заявок. СМО может быть значительно сложней. Выходящие потоки обслуженных заявок, заявок, получивших отказ и заявок, обслуживание которых было по той или иной причине прервано в свою очередь могут образовывать входящие потоки для других подсистем СМО. Так, поток пациентов на прием к терапевту разделяется после обслуживания на выходящий поток заявок, обслуживание которых завершено, и потоки заявок на обслуживание к хирургу, кардиологу и другим специалистам. Источник входящего потока в зависимости от целей исследования может рассматриваться как нечто внешнее по отношению к СМО или же как часть системы.

Насколько для организации массового обслуживания нужна теория? Действительно, во многих случаях руководящие решения опираются на опыт и интуицию. И не всегда эти решения оказываются плохими. Однако обратимся к примеру из третьей главы: пусть пропускная способность городского травматологического пункта – 10 пациентов в час, а в городе в этот период времени случается 9 травм в час. На первый взгляд, здесь все в порядке – травматологический пункт должен справиться с работой. И все же, согласно теории, средняя длина очереди у пункта составит 8,1 человека, а среднее время пребывания пациента в очереди – 0,9 часа (54 минуты). При этом 10 % времени работы пункта придется на простой. Например, если смена продолжается 6 часов, то из них 36 минут персонал простаивает. Другой пример: малое количество касс в супермаркете приведет к росту очередей в определенные периоды работы магазина и потере части покупателей, а большое – к большому простою и неэффективным расходам на заработную плату кассиров. Оптимальное решение опять неочевидно. Мы только что рассмотрели два самых простых примера. А если речь пойдет об организации работы районной поликлиники? Это уже проектирование сложной системы!

В основу предлагаемого учебного пособия лег курс лекций, в течение нескольких лет читаемый автором для студентов, обучающихся по направлению «Информационные системы и технологии». Пособие призвано помочь студенту овладеть первичными навыками исследования СМО и построения их моделей. К сожалению, объем книги не позволил автору в полной мере реализовать свои планы. Однако, учитывая опыт преподавания дисциплины, он стремился тщательно разобрать вопросы, вызывающие у студентов наибольшее затруднение, дать приводимые доказательства достаточно подробно, сделать изложение доступным для широкого круга читателей. Небольшая книга не может дать ответы на широкий спектр вопросов, но иногда, прежде чем открыть солидную монографию, человеку нужно прочитать «тонкую» книгу, чтобы понять, кому и зачем нужна теория, и чтобы возникла заинтересованность.

Пособие состоит их трех глав. Первая посвящена входящим потокам – простейшему и простейшему нестационарному. Во второй рассматриваются уравнения Колмогорова и примеры их решения для достаточно простых СМО. Таким образом, для ряда СМО получены аналитические решения дифференциальных уравнений, описывающих их работу. В более сложных случаях аналитическое решение получить непросто. Однако в широком классе задач вероятности возможных состояний СМО быстро приближаются к некоторым предельным значениям – установившемуся решению. В третьей главе анализируются системы гибели и размножения для случая установившихся решений.

Значительное влияние на структуру и содержание предлагаемого пособия оказали работы А. Я. Хинчина [8], Б. В. Гнеденко и И. Н. Коваленко [3], Л. Г. Лабскера и Л. О. Бабешко [7].

В качестве источника задач для творческого усвоения и закрепления пройденного материала можно рекомендовать книгу О. А. Новикова и С. И. Петухова [6].

Замечания и предложения можно направлять по одному из адресов:
belyi@petrsu.ru *или* ***kurs_ belyi1@mail.ru****.*

Глава 1. Входящий поток

Уже давно в самых разных сферах человеческой деятельности для описания процессов, протекающих во времени, широкое распространение получило понятие **поток событий**, которое означает последовательность событий, разделенных некоторыми интервалами времени. Так, одним из ключевых понятий финансовой математики является поток платежей, автодорожники говорят о потоке машин. В теории массового обслуживания события обычно называют заявками или требованиями и, соответственно, рассматривают потоки заявок или требований.

События, образующие поток, могут происходить как в фиксированные, так и в случайные моменты времени. Например, ваши дни рождения приурочены к фиксированным моментам, а встречи знакомых во время вашей прогулки по городу, скорее всего, образуют поток случайных событий. Часто случайные факторы накладываются на определенную закономерность. Так, даже при наличии четкого расписания прибытия автобусов на станцию реальные моменты этих событий могут отличаться от плановых в результате непредвиденных задержек: пробок на дорогах, остановок у шлагбаума, погодных условий и т. д. Поэтому при построении математических моделей мы можем рассматривать моменты наступления одних и тех же событий как фиксированные или как случайные, в зависимости от цели исследования, от разброса фактических моментов наступления событий относительно ожидаемых, от требований к точности. И все же часто, в силу природы исследуемых явлений, события приходится рассматривать именно как случайные.

Вопрос о том, что в той или иной ситуации можно считать событием, обычно не вызывает затруднений. Если нас интересует работа диспетчера такси, очевидно, что событиями будут звонки клиентов на служебный телефон, но никак не звонки ее подруг на личный мобильный телефон. Другой вопрос – **однородность потока**. Однородность означает возможность привлечь для обработки событий одни и те же технологии, ресурсы системы. На практике часто неоднородный поток заменяют набором однородных потоков. Например, неоднородный поток бытовой техники в ремонтной мастерской удобно разбить на однородные потоки холодильников, чайников, утюгов и т. д. Это разбиение мотивируется спецификой обслуживания соответствующих заявок и различными квалификационными требованиями к мастерам по ремонту техники. Здесь вопрос однородности заявок тесно связан с вопросом глубины специализации персонала мастерской.

Под источником входящего потока часто, но не всегда, подразумевают нечто внешнее по отношению к системе массового обслуживания, свойства которого не зависят от особенностей функционирования системы. Так, входящий поток станции скорой помощи – множество телефонных вызовов врача – не зависит от организации работы станции.

§ 1.1. Определение простейшего потока

В теории массового обслуживания наибольшее распространение получил простейший поток. Это обусловлено тем, что он является достаточно адекватной, математически обоснованной и насколько возможно простой моделью многих реальных потоков.

Простейшим нызывают поток однородных событий, обладающий следующими тремя свойствами:

1. **Стационарностью**. Стационарность потока означает, что для любых вещественных T, $t > 0$ и целого $k \geq 0$ вероятность появления k событий на интервале $(T, T + t)$ не зависит от T.

2. **Отсутствием последействия**. Под отсутствием последействия подразумевают независимость вероятности появления k событий на интервале $(T, T + t)$ от количества и времени появления событий до момента T. В дальнейшем вероятность появления k событий на интервале длины t в простейшем потоке будем обозначать $P_k(t)$.

3. **Ординарностью**. Ординарность потока означает выполнение равенства $P_{>1}(h) = \circ(h)$. Здесь $P_{>1}(h)$ – вероятность появления более одного события за время h, а $\circ(h)$ – произвольная вещественная функция h, бесконечно малая более высокого порядка, чем h. То есть

$$\lim_{h \to 0} \frac{P_{>1}(h)}{h} = 0.$$

Отсюда следует равенство нулю вероятности появления одновременно двух и более событий (разумеется, последнее не исключает возможность появления сразу двух и более событий).

Простейший поток является таким же абстрактным математическим объектом, как прямая линия в геометрии. Стационарность, отсутствие последействия и ординарность не более чем допущения. Так, поток вызовов скорой помощи можно считать стационарным лишь на некоторых ограниченных интервалах времени суток. При небольшом количестве радиоактивного вещества в потоке распадов атомов последействие практически не имеет места, но в случае большой массы того же вещества наблюдается цепная реакция, т. е. вероятность распада некоторого количества атомов на заданном интервале времени зависит от количества произошедших ранее распадов. Допущение об ординарности также периодически нарушается. Например, в железнодорожной кассе иногда приобретаются билеты сразу на целую группу туристов. Увы, мы живем не в идеальном мире, но это не повод отказаться от любых попыток понять его.

§ 1.2. Уравнения простейшего потока

Введем обозначение $P_0(1) = \theta$, где $P_0(t)$ – вероятность того, что за время t не произойдет ни одного события. Разобьем единичный интервал времени на n равных частей. Тогда для того, чтобы на всем интервале не произошло ни одного события, необходимо и достаточно, чтобы ни одного события не произошло на каждом из n частных интервалов. Поскольку $P_0(1/n)$ зависит только от длины интервала, при условии отсутствия последействия $P_0(1) = [P_0(1/n)]^n = \theta$ или $P_0(1/n) = \theta^{\frac{1}{n}}$. Аналогично при натуральном k получим $P_0(k/n) = \theta^{\frac{k}{n}}$. Пусть вещественное число $t > 0$. Тогда для любых t и n можно найти такое k, что

$$\frac{k-1}{n} \leq t \leq \frac{k}{n} \qquad \Longrightarrow$$

$$\Longrightarrow P_0(\frac{k-1}{n}) \geq P_0(t) \geq P_0(\frac{k}{n}) \quad \text{или} \quad \theta^{\frac{k-1}{n}} \geq P_0(t) \geq \theta^{\frac{k}{n}}.$$

Тогда

$$\lim_{n\to\infty} \frac{k}{n} = \lim_{n\to\infty} \frac{k-1}{n} = t \quad \text{и} \quad P_0(t) = \theta^t.$$

Поскольку вероятность всегда принимает значения из интервала $[0;1]$, мы должны рассмотреть три случая: $\theta = 0$, $\theta = 1$ и $0 < \theta < 1$. В первом случае $P_0(t) = 0$ для любого $t > 0$ и, следовательно, на любом сколь угодно малом интервале произойдет бесконечное множество событий, что противоречит принципу ординарности потока. Во втором случае $P_0(t) = 1$ и поток, как таковой, отсутствует. Остается только третий случай. Введем замену переменной $\theta = e^{-\lambda \cdot t}$, где $\lambda > 0$ – некоторый вещественный параметр. Тогда, $\theta \in (0;1] \quad \Longrightarrow \quad \lambda \in [0;+\infty)$. Таким образом, вероятность отсутствия событий на интервале длины t задается равенством $P_0(t) = e^{-\lambda t}$. Смысл параметра λ мы выясним ниже. Поскольку $e^\alpha - 1 = \alpha + \circ(\alpha)$ или $e^\alpha = 1 + \alpha + \circ(\alpha)$, мы можем записать равенство, которое нам в ближайшее время пригодится: $P_0(h) = 1 - \lambda \cdot h + \circ(h)$. Теперь докажем лемму.

Лемма

В простейшем потоке $P_1(h) = \lambda \cdot h + \circ(h)$. То есть вероятность появления одного события за время h с точностью до бесконечно малой более высокого порядка, чем h, пропорциональна h.

Доказательство

Для любого $h \in [0; +\infty]$, $\quad P_0(h) + P_1(h) + P_{>1}(h) = 1$. Подставив в последнее равенство $P_0(h) = 1 - \lambda \cdot h + \circ(h)$ *и* $P_{>1}(h) = \circ(h)$, получим $P_1(h) = \lambda \cdot h + \circ(h)$, что и требовалось доказать. $\square$

Напомним, что $\circ(h)$ – произвольная бесконечно малая величина, большего, чем h, порядка. Поэтому $\circ(h) \pm \circ(h) = \circ(h)$ и $C \cdot \circ(h) = \circ(h)$, где $\forall\ C \in (0; \infty)$ – вещественная константа.

Рассмотрим интервал времени длины $t + h$. Пусть на этом интервале произошло k событий. Тогда, если на участок длины t пришлось j событий, то на участок длины h придется $k - j$ событий, где $j = 0, 1, \ldots, k$. По формуле полной вероятности для $k = 1$ имеет место равенство

$$
\begin{aligned}
P_1(t+h) &= P_1(t) \cdot P_0(h) + P_0(t) \cdot P_1(h) = \\
&= P_1(t) \cdot (1 - \lambda \cdot h + \circ(h)) + P_0(t) \cdot (\lambda \cdot h + \circ(h));
\end{aligned}
$$

$$
P_1(t+h) = P_1(t) \cdot (1 - \lambda \cdot h) + P_0(t) \cdot \lambda \cdot h + \circ(h). \qquad (1)
$$

Аналогично для $k > 1$

$$\begin{aligned} P_k(t+h) &= \sum_{j=0}^{k} P_j(t)\cdot P_{k-j}(h) = \\ &= P_k(t)\cdot P_0(h) + P_{k-1}(t)\cdot P_1(h) + \sum_{j=0}^{k-2} P_j(t)\cdot P_{k-j}(h); \end{aligned}$$

$$0 \le \sum_{j=0}^{k-2} P_j(t)\cdot P_{k-j}(h) \le \sum_{j=0}^{k-2} P_{k-j}(h) = \sum_{j=2}^{k} P_j(h) = P_{>1}(h) = \circ(h).$$

То есть

$$\sum_{j=0}^{k-2} P_j(t)\cdot P_{k-j}(h) = \circ(h).$$

Значит,

$$\begin{aligned} P_k(t+h) &= P_k(t)\cdot P_0(h) + P_{k-1}(t)\cdot P_1(h) + \circ(h) = \\ &= P_k(t)\cdot(1-\lambda\cdot h+\circ(h)) + P_{k-1}(t)\cdot(\lambda\cdot h+\circ(h)) + \\ &+ \circ\,(h); \end{aligned}$$

$$P_k(t+h) = P_k(t)\cdot(1-\lambda\cdot h) + P_{k-1}(t)\cdot\lambda\cdot h + \circ(h). \qquad (2)$$

Тогда из (1) и (2) для $k \ge 1$ следует

$$\frac{P_k(t+h)-P_k(t)}{h} = \lambda\cdot P_{k-1}(t) - \lambda\cdot P_k(t) + \frac{\circ(h)}{h}. \qquad (3)$$

Хотя для $k=0$ уже получено решение $P_0(t) = e^{-\lambda\cdot t}$, составим и для этого случая равенство

$$\frac{P_0(t+h)-P_0(t)}{h} = -\lambda \cdot P_0(t) + \frac{\circ(h)}{h}. \tag{4}$$

Устремив к нулю h в левых и правых частях (3) и (4), получим бесконечную систему дифференциальных уравнений:

$$\begin{cases} P_0'(t) = -\lambda \cdot P_0(t); \\ P_1'(t) = \lambda \cdot P_0(t) - \lambda \cdot P_1(t); \\ \ldots \\ P_k'(t) = \lambda \cdot P_{k-1}(t) - \lambda \cdot P_k(t). \end{cases} \tag{5}$$

Начальные условия:

$$P_0(0) = 1, \quad P_1(0) = \ldots P_k(0) = 0. \tag{6}$$

Первый способ решения системы (5)

Подставим во второе уравнение выражение $P_0(t)$, запишем уравнение в виде

$$P_1'(t) + \lambda \cdot P_1(t) = \lambda \cdot e^{-\lambda \cdot t}$$

и, учтя начальное условие $P_1(0) = 0$, найдем решение:

$$P_1(t) = \lambda \cdot t \cdot e^{-\lambda t}.$$

Теперь, подставив в третье уравнение выражение $P_1(t)$, получим уравнение

$$P_2'(t) + \lambda \cdot P_2(t) = \lambda^2 \cdot t \cdot e^{-\lambda \cdot t}$$

и, учтя начальное условие $P_2(0) = 0$, найдем его решение: $P_2(t) = \frac{(\lambda \cdot t)^2}{2!} e^{-\lambda \cdot t}$. Процесс будем продолжать до тех пор, пока мы не догадаемся, что вероятность появления k событий за время t равна

$$P_k(t) = \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda \cdot t}. \tag{7}$$

Теперь, когда мы знаем результат, нетрудно обосновать его, прибегнув к методу математической индукции. Здесь мы пропустили собственно сам процесс решения уравнений. Первый способ позволяет получить решение, опираясь на минимальный математический аппарат. Однако в некоторых случаях этот способ может оказаться слишком громоздким. Теперь рассмотрим более компактный метод решения бесконечной системы линейных дифференциальных уравнений.

Второй способ решения системы (5)

Введем производящую функцию $\phi(t, z) = \sum_{k=0}^{\infty} P_k(t) \cdot z^k$. Заметим, что из (6) следует $\phi(0, z) = 1$. В уравнениях (5) левые и правые части соответственно индексу при P умножим на z^k и просуммируем отдельно левые и правые части. Тогда

$$\begin{aligned}
\sum_{k=0}^{\infty} P_k'(t) \cdot z^k &= \lambda \cdot \sum_{k=1}^{\infty} P_{k-1}(t) \cdot z^k - \lambda \cdot \sum_{k=0}^{\infty} P_k(t) \cdot z^k; \quad (8) \\
\sum_{k=0}^{\infty} P_k'(t) \cdot z^k &= (\sum_{k=0}^{\infty} P_k(t) \cdot z^k)_t' = \frac{\partial \phi(t, z)}{\partial t}; \\
\sum_{k=1}^{\infty} P_{k-1}(t) \cdot z^k &= z \cdot \sum_{k=1}^{\infty} P_{k-1}(t) \cdot z^{k-1} = z \cdot \phi(t, z).
\end{aligned}$$

Суммы в левой и правой частях (8) можно переписать в виде $\frac{\partial\phi(t,z)}{\partial t} = \lambda\cdot(z-1)\cdot\phi(t,z)$, после чего решение бесконечной системы сводится к решению одного уравнения.

$$\frac{\partial ln(\phi(t,z))}{\partial t} = \lambda\cdot(z-1) \qquad \Longrightarrow \qquad ln(\phi(t,z)) = \lambda\cdot(z-1)\cdot t + ln|C|$$

или

$$\phi(t,z) = C\cdot e^{\lambda\cdot(z-1)\cdot t},$$

где вещественное $C = const$. Учитывая равенство $\phi(0,z) = 1$, получим $C = 1$ и

$$\phi(t,z) = e^{\lambda\cdot(z-1)\cdot t} = e^{\lambda\cdot z\cdot t}\cdot e^{-\lambda\cdot t} = \sum_{k=0}^{\infty}\frac{(\lambda\cdot t)^k}{k!}e^{-\lambda\cdot t}\cdot z^k = \sum_{k=0}^{\infty}P_k(t)\cdot z^k.$$

Осталось только приравнять в последнем равенстве коэффициенты при степенях z. Мы снова получили $P_k(t) = \frac{(\lambda\cdot t)^k}{k!}e^{-\lambda\cdot t}$. Последняя формула в теории вероятностей известна как формула **распределения Пуассона**.

§ 1.3. Свойства простейшего потока

На рис. 1 представлены графики функций (7) при $k = 0, 1, 2, 3$ и $\lambda = 4$. Взяв (при $k > 0$) производную одной из функций (7) и приравняв ее к нулю, легко найдем точку максимума функции $t_{max} = \frac{k}{\lambda}$. Значение вероятности в точке максимума $P_k(t_{max}) = \frac{k^k}{k!}\cdot e^{-k}$. Хотя мы предполагали $k > 0$, формула $t_{max} = \frac{k}{\lambda}$ справедлива и при $k = 0$.

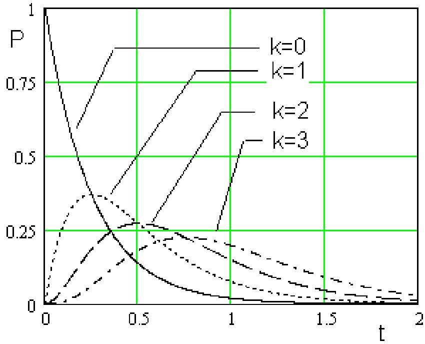

Рис. 1. Распределение Пуассона

Нетрудно также аналитически доказать еще одну отраженную на рисунке закономерность. График каждой функции пересекает график следующей по значению k функции в точке ее максимума.

Как доказано выше, $P_0(t) = e^{-\lambda \cdot t}$ – вероятность того, что за время t не произойдет ни одного события или, иначе говоря, того, что промежуток между двумя событиями больше t.

Тогда $P(\tau < t) = F(t) = 1 - e^{-\lambda \cdot t}$ – вероятность того, что время τ между двумя событиями меньше t, а плотность вероятности

$$f(t) = F'(t) = \lambda \cdot e^{-\lambda \cdot t}.$$

$$\bar{T}_{\textit{ож}} = M(t) = \int_{-\infty}^{+\infty} f(t) \cdot t \cdot dt = \int_{0}^{+\infty} \lambda \cdot e^{-\lambda \cdot t} \cdot t \cdot dt = \frac{1}{\lambda}$$

– математическое ожидание времени t между двумя событиями.

$$M(t^2) = \int_{-\infty}^{+\infty} f(t) \cdot t^2 \cdot dt = \int_{0}^{+\infty} \lambda \cdot e^{-\lambda \cdot t} \cdot t^2 \cdot dt = \frac{2}{\lambda^2}$$

– математическое ожидание t^2 .

$$D(t) = M(t^2) - (M(t))^2 = \frac{1}{\lambda^2}, \qquad \sigma(t) = \sqrt{D(t)} = \frac{1}{\lambda}$$

– соответственно дисперсия и среднее квадратичное отклонение времени между двумя событиями от ожидаемого. Иначе говоря, среднее время между заявками на обслуживание $\frac{1}{\lambda}$ и среднее отклонение от среднего также $\frac{1}{\lambda}$.

Заметим также, что $M(t^k) = \frac{k!}{\lambda^k}$. Теперь определим ожидаемое количество событий за время t и его разброс.

$$\begin{aligned} M(k) &= \sum_{k=0}^{\infty} k \cdot P_k(t) = \sum_{k=1}^{\infty} k \cdot \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda \cdot t} = \\ &= e^{-\lambda \cdot t} \cdot \lambda \cdot t \sum_{k=1}^{\infty} k \cdot \frac{(\lambda \cdot t)^{k-1}}{k!} = e^{-\lambda \cdot t} \cdot t \cdot (e^{\lambda \cdot t})'_t = \\ &= \lambda \cdot t. \end{aligned}$$

Значит, математическое ожидание количества событий за время t равно $\lambda \cdot t$, а λ – ожидаемое количество событий, приходящееся на единицу времени. Величину λ называют интенсивностью входящего потока, или параметром потока.

Аналогично найдем математическое ожидание k^2:

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

дисперсию и среднее квадратичное отклонение k:

$$D(k) = M(k^2) - [M(k)]^2 = \lambda \cdot t, \qquad \sigma(k) = \sqrt{D(k)} = \sqrt{\lambda \cdot t}.$$

Таким образом, ожидаемое количество событий за время t равно $\lambda \cdot t \pm \sqrt{\lambda \cdot t}$.

§ 1.4. Простейший нестационарный поток

Как было отмечено выше, стационарность потока является всего лишь допущением. По большей части реальные потоки не являются стационарными. Например, интенсивность потока отказов технического устройства, как правило, является функцией времени $\lambda(t)$ и на графике (рис. 2) представляет собой известную U-образную кривую. Неисправности приборов случаются чаще всего в начале и в конце срока их эксплуатации. Любопытно, что и человек болеет чаще всего в молодости и в старости. Достоверную статистику по выходу из строя бытовой техники можно найти в соответствующих мастерских. Естественно, статистика отказов оборудования ведется на транспорте, на производстве и в других сферах деятельности человека.

Для нижней части U-образной кривой характерен пологий уча-

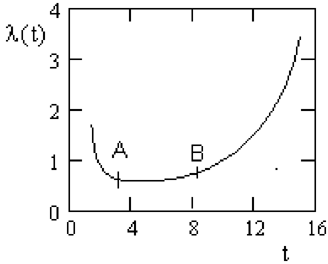

Рис. 2. Кривая отказов технического устройства

сток, на котором интенсивность отказов можно считать постоянной (участок AB на рис. 2).

Теперь вернемся к заголовку параграфа. Поскольку простейший поток мы определили как стационарный, заголовок может показаться противоречивым. Однако в данном случае речь идет всего лишь о новом определении: простейший нестационарный поток – это ординарный поток без последействия. Простейший нестационарный поток также называют пуассоновским потоком.

Поскольку в нестационарном потоке вероятность появления k событий на интервале зависит не только от длины интервала, но и от его начала t_0, в дальнейшем будем обозначать $P_k(t_0, t)$ вероятность появления k событий на интервале (t_0, t) . Добавим к ординарности условие, соответствующее условию леммы из § 1.2.

Тогда

$$P_{>1}(t, t+h) = \circ(h), \qquad P_1(t, t+h) = \lambda(t) \cdot h + \circ(t). \tag{9}$$

Подставим в равенство

$$P_0(t, t+h) + P_1(t, t+h) + P_{>1}(t, t+h) = 1$$

выражения (9) и придем к равенству

$$P_0(t, t+h) = 1 - \lambda(t) \cdot h + \circ(h). \tag{10}$$

Подставив (10) в равенство

$$P_0(t_0, t+h) = P_0(t_0, t) \cdot P_0(t, t+h),$$

после несложных преобразований получим:

$$\frac{P_0(t_0, t+h) - P_0(t_0, t)}{h} = -\lambda(t) \cdot P_0(t_0, t) + \circ(h). \tag{11}$$

Аналогично

$$P_k(t_0, t+h) = \sum_{j=0}^{k} P_j(t_0, t) \cdot P_{k-j}(t, t+h).$$

Повторив рассуждения из § 1.2, получим отношение

$$\frac{P_k(t_0, t+h) - P_k(t_0, t)}{h} = -\lambda(t) \cdot [P_{k-1}(t_0, t) - P_k(t_0, t)] + \circ(h). \tag{12}$$

Устремив в (11) и (12) h к нулю, получим бесконечную систему дифференциальных уравнений:

$$\begin{cases} \frac{\partial P_0(t_0,t)}{\partial t} = -\lambda(t) \cdot P_0(t_0,t); \\ \frac{\partial P_1(t_0,t)}{\partial t} = -\lambda(t)(P_0(t_0,t) - P_1(t_0,t)); \\ \dots \\ \frac{\partial P_k(t_0,t)}{\partial t} = -\lambda(t)(P_{k-1}(t_0,t) - P_k(t_0,t)); \\ \dots \end{cases} \tag{13}$$

Введем вспомогательную функцию $P_{-1}(t_0,t) \equiv 0$. Тогда

$$\forall \ k \geq 0 \qquad \frac{\partial P_k(t_0,t)}{\partial t} = \lambda(t) \cdot [P_{k-1}(t_0,t) - P_k(t_0,t)]. \tag{14}$$

Введем производящую функцию

$$\phi(t_0,t,z) = \sum_{k=0}^{\infty} P_k(t_0,t) \cdot z^k. \tag{15}$$

Умножим левые и правые части уравнений (14) на соответствующие степени z и просуммируем их.

$$\sum_{k=0}^{\infty} \frac{\partial P_k(t_0,t)}{\partial t} \cdot z^k = \lambda(t) \cdot [\sum_{k=0}^{\infty} P_{k-1}(t_0,t) z^k - P_k(t_0,t) z^k],$$

$$\frac{\partial \phi(t_0,t,z)}{\partial t} = \lambda(t) \cdot \phi(t_0,t,z) \cdot (z-1), \qquad \frac{\partial \ln \phi(t_0,t,z)}{\partial t} = \lambda(t) \cdot (z-1).$$

Проинтегрируем последнее равенство по t:

$$\ln(\phi(t_0, t, z)) - \ ln(\phi(t_0, t_0, z)) = (z-1) \cdot \int_{t_0}^{t} \lambda(t) dt.$$

Из начальных условий:
$P_0(t_0, t_0) = 1, P_1(t_0, t_0) = \cdots = P_k(t_0, t_0) = \cdots = 0$ следует $\phi(t_0, t_0, z) = 1$. Введем обозначение $\Lambda(t_0, t) = \int_{t_0}^{t} \lambda(t) \cdot t$. Тогда

$$\begin{aligned} \phi(t_0, t, z) &= e^{(z-1)\Lambda(t_0,t)} e^{z\Lambda(t_0,t)} = \\ &= e^{-\Lambda(t_0,t)} \cdot e^{zcdot\Lambda(t_0,t)} = \\ &= e^{-\Lambda(t_0,t)} \cdot \sum_{k=0}^{\infty} \frac{[\Lambda(t_0, t)]^k}{k!} \cdot z^k. \end{aligned} \tag{16}$$

Приравняв в (15) и (16) коэффициенты при соответствующих степенях z, получим:

$$P_k(t_0, t) = \frac{[\Lambda(t_0, t)]^k}{k!} \cdot e^{-\Lambda(t_0,t)}. \tag{17}$$

$$\text{Пусть} \qquad \bar{\lambda} = \frac{\Lambda(t_0, t)}{t - t_0} = \frac{\int_{t_0}^{t} \lambda(t) dt}{(t - t_0)}$$

– средняя интенсивность входящего потока на интервале (t_0, t), $\tau = t - t_0$ – длина интервала. Тогда $\Lambda(t_0, t) = \bar{\lambda} \cdot \tau$ и уравнение (17) примет вид

$$P_k(t_0, t) = \frac{[\bar{\lambda} \cdot \tau]^k}{k!} \cdot e^{-\bar{\lambda} \cdot \tau}.$$

Таким образом, мы снова пришли к формуле Пуассона.

Глава 2. Марковская модель СМО

Функционирование СМО мы будем рассматривать как **случайный процесс с непрерывным временем и дискретным множеством состояний**. Следовательно, в любой момент времени $t \in [0; +\infty)$ система находится в одном состоянии из заданного конечного или счетного набора.

Например, в цехе имеется десять однотипных станков. Станки обслуживает один мастер по ремонту. Таким образом, соответствующая система может находиться в одном из 11 состояний: S_0 – все станки исправны, S_1 – один в ремонте и девять в рабочем состоянии, S_2 – один в ремонте, один в очереди на ремонт и восемь работают, ..., S_{10} – один в ремонте и девять в очереди. Очевидно, момент отказа станка и время, необходимое для устранения неисправности, – случайные величины. В процессе функционирования система иногда переходит из одного состояния в другое. Более того, теоретически в любой момент времени система может находиться в любом из перечисленных выше состояний. Поэтому имеет смысл говорить только о вероятностях соответствующих состояний: $P_0(t), P_1(t), P_2(t) \dots P_{10}(t)$.

Случайный процесс называется **марковским**, если для любого момента времени t условные вероятности всех состояний системы в будущем зависят только от состояния системы в момент t и не зависят от того, когда и каким образом она пришла в это состояние. Иначе говоря, будущее зависит от прошлого только через настоящее. Марковский процесс называют также процессом без последействия. Мы будем считать процесс функционирования СМО марковским. Хотя марковская модель СМО не является единственно возможной, она достаточно адекватно отражает широкий класс реальных систем.

Пусть система в любой момент времени может находиться в одном из n возможных состояний S , где $i = 1, 2, \ldots n$. В частности, не исключается случай $n = \infty$. То есть множество исходов не более чем счетно. Иногда нам будет удобней говорить не «состояние S_i », а «i-е состояние». Нумерацию состояний мы часто будем начинать не с единицы, а с нуля. Сделаем допущение, что вероятность перехода системы за время h из $i-$ в $j-$ состояние задается равенством

$$P_{ij}(h) = \lambda_{ij} \cdot h + \circ(h), \textit{где } i \neq j. \tag{18}$$

То есть на небольшом интервале времени вероятность перехода системы из i-го в j-е состояние пропорциональна длине интервала. Равенства (18), очевидно, делают процесс марковским. Величину λ_{ij} , где $i \neq j$, назовем **интенсивностью перехода** из i-го в j-е состояние. В общем случае λ_{ij} могут зависеть от времени, но здесь мы ограничимся случаем постоянных интенсивностей.

Равенства (18) аналогичны равенству, доказанному в первой главе для простейшего потока $P_1(h) = \lambda \cdot h + \circ(h)$. Более того, сам поток однородных событий можно интерпретировать как случайный процесс накопления событий. Пусть $k(t)$ – число событий, произошедших до момента t. Каждая реализация такого случайного процесса представляет собой ступенчатую функцию (рис. 3), значение которой увеличивается на единицу с появлением очередного события.

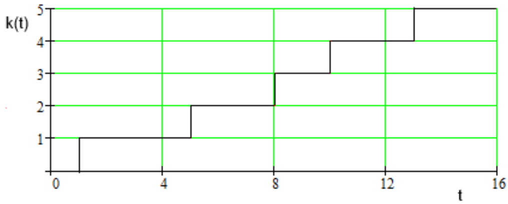

Рис. 3. Простейший поток как случайный процесс

Можно также связать описанный случайный процесс с системой, имеющей множество состояний S_i, где $i = 0, 1, 2, \ldots, n$. В данном случае i – количество событий. Тогда интенсивности переходов:

$$\lambda_{ij} = \begin{cases} \lambda, \ \textit{если } j = i + 1, \ \textit{где } i, j = 0, 1 \ldots \infty; \\ 0, \textit{иначе.} \end{cases}$$

Здесь под λ мы, как и в первой главе, подразумеваем интенсивность входящего потока событий. В данном случае на множестве состояний $S_0, S_1, S_2, \ldots$ допустимы только переходы слева направо в порядке возрастания номеров.

СМО с дискретным множеством состояний мы часто будем схематически представлять в виде направленного графа, вершинами которого являются состояния, а дугами – допустимые переходы из одного состояния в другое.

§ 2.1. Уравнения Колмогорова

Пусть состояния СМО занумерованы натуральными числами $i = 1, 2, \ldots n$. Заметим, что вероятность (18) перехода $P_{ik}(h) = P(S_k(t+h)/S_i(t))$ – условная вероятность, иначе говоря, вероятность того, что система в момент времени $t+h$ оказалась в состоянии S_k при условии, что в момент t система находилась в состоянии S_i. Разумеется, $k = 1, 2, \ldots n$.

Если здесь t настоящее, то, таким образом, вероятности всех возможных состояний в будущем зависят только от состояния в настоящем. Обозначим также

$$P_{kk}(h) = P(S_k(t+h)/S_k(t))$$

– вероятность того, что система за время h не изменит текущее состояние. Поскольку находящаяся в состоянии S_k система за время h либо перейдет в какое-либо иное состояние, либо останется в S_k,

$$\sum_{i=1}^{n} P_{ki}(h) = 1, \text{ откуда } P_{kk}(h) = 1 - \sum_{i \neq k} P_{ki}(h).$$

По формуле полной вероятности

$$P_k(t+h) = \sum_{i \neq k} P_i(t) \cdot P_{ik}(h) + P_k(t) \cdot (1 - \sum_{i \neq k} P_{ki}(h)).$$

Подставим в последнюю формулу значения из (18), получим:

$$\begin{aligned} P_k(t+h) &= \sum_{i\neq k} P_i(t)(\lambda_{ik}\cdot h + \circ(h)) + \\ &+ P_k(t)\cdot(1-\sum_{i\neq k}(\lambda_{ki}\cdot h+\circ(h))) = \\ &= \sum_{i\neq k}\lambda_{ik}\cdot h\cdot P_i(t) + P_k(t)\cdot(1-\sum_{i\neq k}\lambda_{ki}\cdot h) + \circ(h). \end{aligned}$$

Отсюда $$\frac{P_k(t+h)-P_j(t)}{h} = \sum_{i\neq k}\lambda_{ij}\cdot P_i(t) - \sum_{i\neq k}\lambda_{ki}P_k(t).$$

Устремив h к нулю, получим линейное дифференциальное уравнение, соответствующее k-му состоянию системы:

$$P_k'(t) = \sum_{i\neq k}\lambda_{ij}\cdot P_i(t) - \sum_{i\neq k}\lambda_{ki}P_k(t). \tag{19}$$

Уравнения (19) для всех состояний СМО образуют систему уравнений Колмогорова, описывающую работу произвольной СМО с постоянными интенсивностями переходов:

$$\begin{cases} P_1'(t) = -\sum_{i\neq 1}\lambda_{1i}\cdot P_1(t) + \lambda_{21}\cdot P_2(t) + \cdots + \lambda_{n1}\cdot P_n(t); \\ P_2'(t) = \lambda_{12}\cdot P_1(t) - \sum_{i\neq 2}\lambda_{2i}\cdot P_2(t) + \cdots + \lambda_{n2}\cdot P_n(t); \\ \ldots \\ P_n'(t) = \lambda_{1n}\cdot P_1(t) + \lambda_{2n}\cdot P_2(t) + \cdots - \sum_{i\neq n}\lambda_{ni}\cdot P_n(t). \end{cases}$$

Начальные условия обычно имеют вид $P_1(0) = 1$, $P_2(0) = P_3(0) = = \cdots = P_n(0)$, поскольку в начале работы СМО находится в некотором исходном состоянии.

Рассмотренной выше СМО соответствует полный направленный граф, т. е. допускаются переходы из любого состояния в любое другое. Граф реальной СМО, скорее всего, окажется неполным. Однако для нас это будет означать только то, что в системе уравнений (19) некоторые интенсивности переходов следует приравнять к нулю.

§ 2.2. Одноканальная СМО с отказами

Самая простая СМО – одноканальная с отказами (рис. 4).

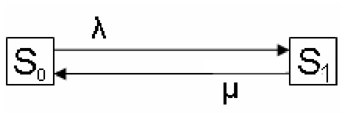

Рис. 4. Одноканальная СМО с отказами

Эта система в любой момент времени может находиться в одном из двух состояний. Состояние S_0 – единственный канал свободен, S_1 – канал занят обслуживанием заявки. Если в момент поступления очередной заявки канал занят, заявка получает отказ, т. е. теряется. Как видно на схеме, интенсивности переходов $\lambda_{01} = \lambda$ и $\lambda_{10} = \mu$. Здесь λ – интенсивность входящего потока, μ – интенсивность потока обслуживания. Предполагается, что время обслуживания – случайная величина с экспоненциальной плотностью распределения $f(t) = \mu \cdot e^{-\mu \cdot t}$. То есть время обслуживания распределено по тому же закону, что и время между двумя соседними заявками. Таким образом, для вероятностей переходов выполняются равенства (18):

$$P_{01}(h) = \lambda \cdot h + \circ(h) \text{ и } P_{10}(h) = \mu \cdot h + \circ(h).$$

Математическое ожидание времени обслуживания

$$\bar{T}_{обсл} = M(t) = \int_{-\infty}^{+\infty} f(t) \cdot t \cdot dt = \int_{0}^{+\infty} \mu \cdot e^{-\mu \cdot t} \cdot t \cdot dt = \frac{1}{\mu}.$$

Это соответствует полученному в первой главе результату для времени ожидания очередной заявки $\bar{T}_{обсл} = \frac{1}{\lambda}$. Естественно возникает вопрос: «Не несет ли в себе параметр μ смысл, аналогичный смыслу интенсивности простейшего потока λ?». Действительно, μ можно определить как ожидаемое количество обслуживаемых в единицу времени заявок при условии, что канал обслуживания работает непрерывно. На самом деле канал обслуживания иногда простаивает, и потому μ не совпадает с интенсивностью выходящего потока, т. е. потока обслуженных заявок.

Запишем систему уравнений Колмогорова для одноканальной СМО с отказами:

$$\begin{cases} P_0'(t) = -\lambda \cdot P_0(t) + \mu \cdot P_1(t); \\ P_1'(t) = \lambda \cdot P_0(t) - \mu \cdot P_1(t). \end{cases}$$

Начальные условия $P_0(0) = 1$ и $P_1(0) = 0$, т. е. в начале работы система готова принять заявку. Подставив в первое уравнение

$$P_1(t) = 1 - P_0(t), \text{ получим } P_0'(t) + (\lambda + \mu) \cdot P_0(t) = \mu. \qquad (20)$$

Решение

1. Найдем решение соответствующего однородного уравнения:

$$P_0'(t) + (\lambda + \mu) \cdot P_0(t) = 0.$$

$$\frac{dP_0}{P_0} = -(\lambda + \mu) \cdot dt = 0 \implies lnP_0(t) = -(\lambda + \mu) \cdot t + ln \mid C \mid,$$

где $C - const$. Тогда $P_0(t) = C \cdot e^{-(\lambda+\mu)\cdot t}$.

2. Найдем одно частное решение исходного уравнения в виде $\rho(t) = \alpha$ методом неопределенных коэффициентов. Подставив $\rho(t) = \alpha$ в (20), получим $(\lambda+\mu)\cdot\alpha = \mu$. Таким образом, $\rho(t) = \frac{\mu}{\lambda+\mu}$.

3. Общее решение неоднородного линейного уравнения складывается из общего решения соответствующего однородного и произвольного частного решения. Следовательно, общее решение уравнения (20) будет иметь вид

$$P_0(t) = C \cdot e^{-(\lambda+\mu)\cdot t} + \frac{\mu}{\lambda + \mu}.$$

Из условия $P_0(0) = 1$ вытекает $C = \frac{\lambda}{\lambda+\mu}$ и искомое решение

$$P_0(t) = \frac{1}{\lambda + \mu} \cdot (\mu + \lambda \cdot e^{-(\lambda+\mu)\cdot t}).$$

$P_1(t)$ найдем как $1 - P_0(t)$ и представим результат в виде

$$\begin{cases} P_0(t) = \frac{1}{\lambda+\mu} \cdot (\mu + \lambda \cdot e^{-(\lambda+\mu)\cdot t}); \\ P_1(t) = \frac{\lambda}{\lambda+\mu} \cdot (1 - e^{-(\lambda+\mu)\cdot t}). \end{cases} \tag{21}$$

Заметим, что $\lim\limits_{t\to+\infty} P_0(t) = \dfrac{\mu}{\lambda+\mu}, \lim\limits_{t\to+\infty} P_1(t) = \dfrac{\lambda}{\lambda+\mu}$.

Таким образом, графики $P_0(t)$ и $P_1(t)$ на бесконечности стремятся к некоторым асимптотам.

Значения $P_0(t) = \frac{\mu}{\lambda+\mu}$ и $P_1(t) = \frac{\lambda}{\lambda+\mu}$ называют **установившимися решениями**, а также **предельными вероятностями**, или **стационарными вероятностями**. В установившихся решениях после P_k мы не пишем в скобках t.
На рис. 5 представлены графики вероятностей состояний системы на временном интервале $[0; 2]$. Как видно, графики очень быстро сливаются с асимптотами. В таких случаях часто сосредотачивают внимание на установившихся решениях. И все же иногда, например, когда речь идет о запуске космического корабля, крайне важно поведение системы именно на начальном временном интервале. На графиках представлены три решения при различных отношениях между интенсивностью входящего потока заявок и интенсивностью обслуживания и, соответственно, три варианта установившегося решения:

1. $\lambda < \mu$ – система чаще свободна, чем занята обслуживанием заявок;

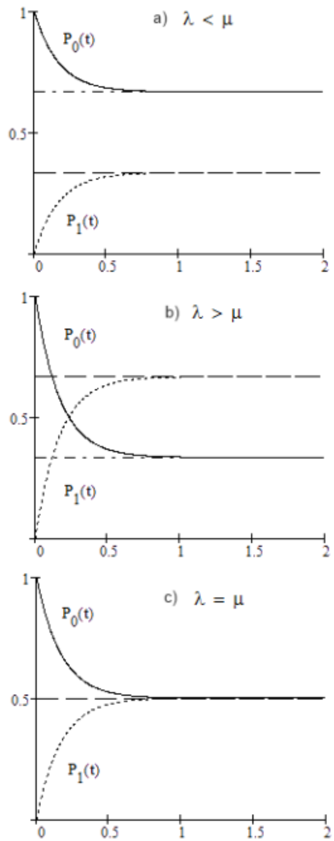

Рис. 5. Одноканальная СМО с отказами: a) $\lambda = 4$ и $\mu = 2$, b) $\lambda = 2$ и $\mu = 4$, c) $\lambda = 3$ и $\mu = 3$

2. $\lambda > \mu$ – система чаще занята;

3. $\lambda = \mu$ – система простаивает ровно в половине случаев.

Установившиеся решения можно получать и непосредственно из уравнений Колмогорова. Для этого достаточно в одном из уравнений (19) заменить переменные $P_k(t)$ на константы P_k и добавить условие $\sum_k P_k = 1$.

Для рассмотренной в этом параграфе СМО система уравнений примет вид

$$\begin{cases} -\lambda \cdot P_0(t) + \mu P_1 = 0; \\ P_0(t) + P_1 = 1. \end{cases} \tag{22}$$

Разумеется, ее решение совпадет с результатом, полученным выше путем предельного перехода.

Характеристики одноканальной СМО с отказами

1. Ожидаемое **время между двумя последовательными заявками**

$$\bar{T}_{Ожид} = \frac{1}{\lambda}.$$

2. Ожидаемое **время обслуживания заявки**

$$\bar{T}_{Обсл} = \frac{1}{\mu}.$$

3. **Относительная пропускная способность** $Q(t) = P_0(t)$ – доля обслуженных заявок в общем количестве поступивших. В данном случае эта величина совпадает с вероятностью того, что единственный канал обслуживания в момент t свободен.

   В пределе $$Q = P_0 = \frac{\mu}{\lambda + \mu}.$$

4. **Абсолютная пропускная способность** $A(t) = \lambda \cdot Q(t) = \lambda \cdot P_0(t)$ – среднее число обслуживаемых в единицу времени заявок.

В пределе $$A = \lambda \cdot P_0 = \frac{\lambda \cdot \mu}{\lambda + \mu}.$$

Поскольку каждая принятая заявка будет обслужена, эта величина здесь совпадает с интенсивностью выходящего потока.

5. **Ожидаемая доля необслуженных заявок** среди поступивших в момент t: $P_{отк}(t) = P_1(t) = 1 - Q(t)$.

В пределе $$P_{отк} = \frac{\lambda}{\lambda + \mu}.$$

Обратим внимание на тот факт, что рассмотренная в этом параграфе система имеет два выходящих потока заявок. В предельном случае:

$$A = \lambda \cdot P_0 = \frac{\lambda \cdot \mu}{\lambda + \mu}$$

– интенсивность потока обслуженных заявок и

$$A = \lambda \cdot P_1 = \frac{\lambda^2}{\lambda + \mu}$$

– интенсивность потока потерянных заявок.

§ 2.3. Дублированная СМО с восстановлением

Теперь рассмотрим одну классическую задачу теории надежности. Некоторое устройство в процессе работы может выходить из строя. Имеется резервное устройство, которое в случае неисправности основного автоматически включается в работу. В этот же момент начинается восстановление основного. Будем считать, что резерв ненагруженный, т. е. во время работы основного устройства резервное не может потерять работоспособность.

Пусть λ – интенсивность потока отказов, μ – интенсивность восстановления. Тогда $\frac{1}{\lambda} = \bar{T}_{отк}$ – ожидаемая наработка на отказ, т. е. среднее время работы устройства до его отказа, $\frac{1}{\mu} = \bar{T}_{восст}$ – ожидаемое время восстановления неисправного устройства, т. е. среднее время устранения неисправности.

Изначально система находится в состоянии S_0 – работает основное устройство. В случае выхода из строя основного устройства, система переходит в состояние S_1 – работает резервное устройство. Если во время работы резервного устройства было восстановлено основное, система возвращается в S_0. Если же до восстановления основного устройства вышло из строя резервное, система переходит в состояние S_2, что фактически означает прекращение работы системы.

Составим по изображенной на рис. 6 схеме систему уравнений

Колмогорова:

$$\begin{cases} P_0'(t) = -\lambda \cdot P_0(t) + \mu \cdot P_1(t); \\ P_1'(t) = \lambda \cdot P_0(t) - (\lambda + \mu) \cdot P_1(t); \\ P_2'(t) = \lambda \cdot P_1(t). \end{cases} \tag{23}$$

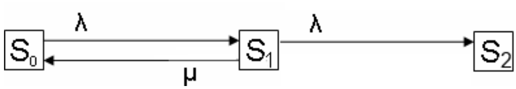

Рис. 6. Дублированная СМО с восстановлением

Начальные условия: $P_0(0) = 1$ и $P_1(0) = P_2(0) = 0$. Из второго уравнения выразим

$$\lambda \cdot P_0(t) = P_1'(t) + (\lambda + \mu) \cdot P_1(t). \tag{24}$$

Левую и правую части первого из уравнений (23) умножим на λ и подставим в полученное уравнение значение $\lambda \cdot P_0(t)$ из (24). После приведения подобных членов получим линейное однородное уравнение второго порядка

$$P_1''(t) + (2 \cdot \lambda + \mu) \cdot P_1'(t) + \lambda^2 \cdot P_1(t) = 0.$$

Составим соответствующий ему характеристический многочлен:

$$k^2 + (2 \cdot \lambda + \mu) \cdot k + \lambda^2$$

и найдем его корни:

$$\begin{aligned} D &= 4\lambda\mu + \mu^2; \\ k_1 &= \frac{-(2\cdot\lambda+\mu) - \sqrt{4\lambda\mu+\mu^2}}{2} < 0; \\ k_2 &= \frac{-(2\cdot\lambda+\mu) + \sqrt{4\lambda\mu+\mu^2}}{2}. \end{aligned}$$

Поскольку $4\lambda\mu+\mu^2 < (2\cdot\lambda+\mu)^2$, k_2 также меньше нуля. Все корни характеристического уравнения отрицательны. Многочлен, вещественные части всех корней которого отрицательны, называют textbfустойчивым. Устойчивость означает, что, как и в примере предыдущего параграфа, все экспоненты, линейной комбинацией которых является решение уравнения, при $t \to +\infty$ стремятся к нулю и уравнение имеет предельное решение.

Общее решение уравнения запишем в виде

$$P_1(t) = e^{-\frac{2\lambda+\mu}{2}\cdot t} \cdot (C_1 \cdot e^{\frac{\sqrt{4\lambda\mu+\mu^2}}{2}\cdot t} + C_2 \cdot e^{-\frac{\sqrt{4\lambda\mu+\mu^2}}{2}\cdot t}).$$

Подставив в уравнение $t = 0$ и применив начальное условие $P_1(0) = 0$, получим $C_2 = -C_1$. С целью экономии пространства и времени введем обозначения:

$$\frac{2\lambda+\mu}{2} = a \ \text{и} \ \frac{\sqrt{4\lambda\mu+\mu^2}}{2} = b.$$

Теперь уравнение перепишем в виде

$$P_1(t) = C \cdot e^{-at} \cdot (e^{bt} - e^{-bt})$$

и возьмем производную

$$P_1'(t) = C \cdot e^{-at} \cdot ((b-a) \cdot e^{bt} + (a+b)e^{-bt}). \qquad (25)$$

Подставив во второе уравнение (23) $t = 0$, и, учитывая $P_0(0) = = 1$ и $P_1(0) = 0$, получим недостающее для дифференциального уравнения второго порядка начальное условие $P_1'(0) = \lambda$. При $t = 0$ из (25) следует: $C \cdot (b - a + a + b) = \lambda$ или $C = \frac{\lambda}{2b}$.

$$P_1(t) = \frac{\lambda}{2b} \cdot e^{-at} \cdot (e^{bt} - e^{-bt}).$$

Теперь легко из (24) получим уравнение

$$P_0(t) = \frac{1}{2b} \cdot e^{-at} \cdot ((b + \frac{\mu}{2}) \cdot e^{bt} + (b - \frac{\mu}{2}) \cdot e^{-bt}).$$

Вероятность безотказной работы системы в течение времени t

$$R(t) = P_0(t) + P_1(t) = \frac{1}{2b} \cdot e^{-at} \cdot ((a+b) \cdot e^{bt} - (a-b) \cdot e^{-bt}). \qquad (26)$$

Разумеется, $P_2(t) = 1 - R(t)$. Причем $\lim_{t\to+\infty} P_2(t) = 1$, и, значит, система в конце концов закончит свой путь в состоянии S_2. Однако нас в этой задаче прежде всего интересует вероятность безотказной работы системы в течение заданного времени. Подставив в правую часть (26) значения a и b, после ряда преобразований получим:

$$R(t) = e^{-\frac{2\lambda+\mu}{2}\cdot t} \cdot [\frac{2\lambda+\mu}{\sqrt{4\lambda\mu+\mu^2}} \cdot sh(\frac{\sqrt{4\lambda\mu+\mu^2}}{2}) + ch(\frac{\sqrt{4\lambda\mu+\mu^2}}{2} \cdot t)]. \qquad (27)$$

Чтобы найти вероятность безотказной работы соответствующей СМО с резервом без восстановления, достаточно устремить к нулю μ.

$$
\begin{aligned}
\lim_{\mu\to 0} e^{-\frac{2\lambda+\mu}{2}} &= e^{-\lambda\cdot t};\\
\lim_{\mu\to 0} ch(\frac{\sqrt{4\lambda\mu+\mu^2}}{2}\cdot t) &= 1;\\
\lim_{\mu\to 0} sh(\frac{\sqrt{4\lambda\mu+\mu^2}}{2})/\sqrt{4\lambda\mu+\mu^2} &= \lim_{b\to 0}\frac{sh(b\cdot t)}{2b} =\\
=\lim_{b\to 0}\frac{e^{bt}-e^{-bt}}{4b} &= \lim_{b\to 0}\frac{t\cdot(e^{bt}-e^{-bt})}{4} = \frac{t}{2}.
\end{aligned}
$$

Следовательно,

$$\lim_{\mu\to 0} R(t) = e^{-\lambda\cdot t}\cdot(1+\lambda\cdot t).$$

На рис.7 представлены графики функции $R(t)$ при $\mu > 0$ (си-

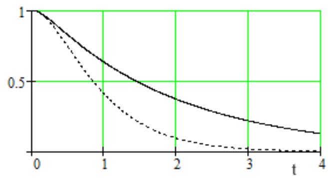

Рис. 7. График R(t) при $\lambda = 2$. Сплошная линия: $\mu = 4$, пунктир: $\mu = 0$

стема с восстановлением) и $\mu = 0$ (система без восстановления). Теперь найдем ожидаемое время наработки системы на отказ. Функция распределения времени безотказной работы $F(t) = 1 - R(t)$. Соответственно, плотность распределения $f(t) = F'(t)$. Вернемся к более компактной, чем (28), формуле (26). Тогда

$$f(t) = \frac{\lambda^2}{2b} \cdot (e^{-(a-b)\cdot t} - e^{-(a+b)\cdot t}).$$

Обозначим наработку на отказ системы $\bar{T}_{сист}$ в отличие от наработки на отказ каждого устройства $\bar{T}_{отк}$:

$$\bar{T}_{сист} = M(t) = \int_0^\infty t \cdot f(t) \cdot dt = \frac{\lambda^2}{2b} \cdot (\frac{1}{(a-b)^2} - \frac{1}{(a+b)^2}).$$

Подставив значения a и b, получим:

$$\bar{T}_{сист} = \frac{2}{\lambda} + \frac{\mu}{\lambda^2}.$$

Здесь $\frac{\mu}{\lambda^2}$ – увеличение наработки системы на отказ за счет восстановления. При $\mu = 0$ значение $\bar{T}_{сист} = \frac{2}{\lambda} = 2 \cdot \bar{T}_{отк}$. То есть наработка на отказ системы без восстановления равна сумме наработок на отказ основного и резервного устройств. Тот же результат можно получить, взяв

$$\begin{aligned}
R(t) &= e^{-\lambda \cdot t} \cdot (1 + \lambda \cdot t), \\
F(t) &= 1 - R(t) \implies \\
f(t) &= F'(t) = \lambda^2 \cdot t \cdot e^{-\lambda \cdot t}. \\
\bar{T}_{сист} &= M(t) = \int_0^\infty t \cdot f(t) \cdot dt = \int_0^\infty \lambda^2 \cdot t^2 \cdot e^{-\lambda \cdot t} \cdot dt = \frac{2}{\lambda}.
\end{aligned}$$

Например, система, где $\lambda = 0,5$ без восстановления будет иметь наработку на отказ $\bar{T}_{сист} = 4$, а при интенсивности восстановления $\mu = 2$ получим $\bar{T}_{сист} = 12$.

§ 2.4. СМО с приоритетными заявками

Системы массового обслуживания с приоритетами мы наблюдаем в железнодорожных кассовых залах, когда вне очереди оформляют билеты ветеранам войн и другим категориям граждан, на которые распространяется соответствующая льгота; в стоматологическом кабинете, где принимают без очереди пациентов с острой болью. Можно привести много подобных примеров.

Системы с приоритетами классифицируют прежде всего по количеству категорий заявок. Так, в военно-полевой медицине принято делить раненых на четыре группы по срочности оказания медицинской помощи. Такая классификация впервые была предложена выдающимся российским хирургом Николаем Ивановичем Пироговым. На телеграфе когда-то выделяли три категории телеграмм: простые, срочные и молния. СМО с приоритетами может быть с отказами или с очередями. Кроме того, при поступлении приоритетной заявки обслуживание «рядовой» может прерываться или же система будет ждать завершения обслуживания. Например, в противовоздушной обороне при появлении более опасных целей система может отпустить неприоритетную заявку и переключиться на обслуживание «дорогих гостей».

Мы рассмотрим СМО (рис. 8) с отказами и с двумя входящими потоками заявок: обычный с интенсивностью λ_1 и приоритетный с интенсивностью λ_2. Интенсивности обслуживания соответствующих заявок – μ_1 и μ_2. Система может находиться в

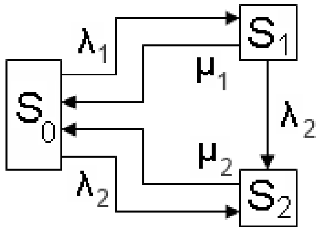

Рис. 8. СМО с приоритетами

трех состояниях: S_0 – свободна, S_1 – обработка обычной заявки, S_2 – обработка приоритетной заявки. Первоначально система находится в состоянии S_0. В случае поступления обычной заявки система переходит в состояние S_1. Если до завершения обслуживания обычной заявки поступила приоритетная, система прерывает обслуживание текущей заявки и приступает к обслуживанию приоритетной, т. е. переходит в состояние S_2. После завершения обслуживания любой заявки система возвращается в исходное состояние S_0.

Обыкновенная заявка получает отказ, если система занята обслуживанием любой другой заявки, приоритетная – только тогда, когда СМО занята обслуживанием другой приоритетной заявки.

Такая система будет иметь пять выходящих потоков, которые соответственно составляют обслуженные приоритетные и обычные заявки, приоритетные и обычные заявки, получившие отказ в обслуживании и, наконец, обычные заявки, принятые на обслуживание, но не обслуженные по вине приоритетных.

Составим систему уравнений Колмогорова:

$$\begin{cases} P_0'(t) = (\lambda_1 + \lambda_2) \cdot P_0(t) + \mu_1 \cdot P_1(t) + \mu_2 \cdot P_2(t); \\ P_1'(t) = \lambda_1 \cdot P_0(t) - (\mu + \lambda_2) \cdot P_1(t); \\ P_2'(t) = \lambda_2 \cdot P_0(t) + \lambda_2 \cdot P_1(t) - \mu_2 \cdot P_2(t). \end{cases} \tag{28}$$

Начальные условия по-прежнему $P_0(0) = 1$ и $P_1(0) = P_2(0) = 0$. Будем искать частные решения системы (29) в виде

$$P_0(t) = \alpha \cdot e^{-kt};$$
$$P_1(t) = \beta \cdot e^{-kt};$$
$$P_2(t) = \gamma \cdot e^{-kt}. \tag{29}$$

В таком случае k должно быть корнем характеристического уравнения

$$\begin{vmatrix} k + \lambda_1 + \lambda_2 & -\mu_1 & -\mu_2 \\ -\lambda_1 & k + \lambda_2 + \mu_1 & 0 \\ -\lambda_2 & -\lambda_2 & k + \mu_2 \end{vmatrix} = 0.$$

Преобразовав левую часть последнего уравнения, разложим на множители полученный многочлен:

$$k \cdot (k + \mu_2 + \lambda_2) \cdot (k + \mu_1 + \lambda_1 + \lambda_2).$$

Итак, характеристический многочлен имеет три различных вещественных корня:

$$k_1 = 0, \qquad k_2 = -(\lambda_2 + \mu_2) \quad \text{и} \quad k_3 = -(\lambda_1 + \lambda_2 + \mu_1).$$

Для каждого корня k подставим (29) в (28) и выберем два первых уравнения из трех линейно зависимых. Найдем решения полученных систем с точностью до постоянного множителя.

1. $k_1 = 0$:

$$\begin{cases} (\lambda_1 + \lambda_2)\alpha - \mu_1 \cdot \beta = \mu_2 \cdot \gamma; \\ -\lambda_1 \cdot \alpha + (\mu_1 + \lambda_2) \cdot \beta = 0; \end{cases} \implies \begin{cases} \alpha = \mu_2 \cdot (\lambda_2 + \mu_1); \\ \beta = \lambda_1 \cdot \mu_1; \\ \gamma = \lambda_2 \cdot (\lambda_1 + \lambda_2 + \mu_1). \end{cases}$$

2. $k_2 = -\lambda_2 - \mu_2$:

$$\begin{cases} (\lambda_1 - \mu_2)\alpha - \mu_1 \cdot \beta = \mu_2 \cdot \gamma; \\ -\lambda_1 \cdot \alpha + (\mu_1 - \mu_2) \cdot \beta = 0; \end{cases} \implies \begin{cases} \alpha = \mu_1 - \mu_2; \\ \beta = \lambda_1; \\ \gamma = -\lambda_1 - \mu_1 + \mu_2. \end{cases}$$

3. $k_3 = -\lambda_1 - \lambda_2 - \mu_1$:

$$\begin{cases} \mu_1 \cdot \alpha - \mu_1 \cdot \beta = \mu_2 \cdot \gamma; \\ -\lambda_1 \cdot \alpha + (\mu_1 - \mu_2) \cdot \beta = 0; \end{cases} \implies \begin{cases} \alpha = 1; \\ \beta = -1; \\ \gamma = 0. \end{cases}$$

Тогда общее решение системы (28) примет вид

$$\begin{pmatrix} P_0(t) \\ P_1(t) \\ P_2(t) \end{pmatrix} = C_1 \cdot \begin{pmatrix} \mu_2 \cdot (\lambda_2 + \mu_1) \\ \lambda \cdot \mu_2 \\ \lambda_2 \cdot (\lambda_1 + \lambda_2 + \mu_1) \end{pmatrix} +$$

$$+ C_2 \cdot e^{-(\lambda_2+\mu_2)\cdot t} \cdot \begin{pmatrix} \mu_1 - \mu_2 \\ \lambda_1 \\ -\lambda_1 - \mu_1 + \mu_2 \end{pmatrix} +$$

$$+ C_3 \cdot e^{-(\lambda_1+\lambda_2+\mu_1)t} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \tag{30}$$

где C_1, C_2 и C_3 – произвольные вещественные константы. Для нахождения частного решения (28), удовлетворяющего начальным условиям, подставим в общее решение $t = 0$.

$$\begin{cases} \mu_2 \cdot (\lambda_2 + \mu_1) \cdot C_1 + (\mu_1 - \mu_2) \cdot C_2 + C_3 = 1; \\ \lambda_1 \cdot \mu_2 \cdot C_1 + \lambda_1 \cdot C_2 - C_3 = 0; \\ \lambda_2 \cdot (\lambda_1 + \lambda_2 + \mu_1) \cdot C_1 + (-\lambda_1 - \mu_1 + \mu_2) \cdot C_2 = 0. \end{cases}$$

Решим систему относительно неизвестных:

$$\begin{cases} C_1 = \frac{1}{(\lambda_1+\lambda_2+\mu_1)\cdot(\lambda_2+\mu_2)}; \\ C_2 = \frac{\lambda_2}{(\lambda_1+\mu_1-\mu_2)\cdot(\lambda_2+\mu_2)}; \\ C_3 = \frac{\lambda_1\cdot(\lambda_1+\lambda_2+\mu_1-\mu_2)}{(\lambda_1+\lambda_2+\mu_1)\cdot(\lambda_1+\mu_1-\mu_2)}. \end{cases}$$

Подставив значения C_1, C_2 и C_3 в (30), получим искомое решение. К сожалению, запись уравнений оказалась слишком громоздкой, но тем не менее мы получили аналитическое решение путем ряда довольно стандартных, рутинных операций, которые любой человек, знакомый с основами теории дифференциальных уравнений, легко может проделать. Аналитическое решение открывает нам большие возможности теоретического исследования различных режимов работы системы.

Непосредственно из (30) путем предельного перехода найдем установившееся решение:

$$\begin{cases} P_0 = \frac{\mu_2 \cdot (\lambda_2 + \mu_1)}{(\lambda_1 + \lambda_2 + \mu_1) \cdot (\lambda_2 + \mu_2)}; \\ P_1 = \frac{\lambda_1 \cdot \mu_2}{(\lambda_1 + \lambda_2 + \mu_1) \cdot (\lambda_2 + \mu_2)}; \\ P_2 = \frac{\lambda_2}{\lambda_2 + \mu_2}. \end{cases} \tag{31}$$

Запись решения системы уравнений (28), как и промежуточные выкладки, значительно упрощается, если взять равные интенсивности обслуживания обычных и привилегированных заявок. На рис. 9 представлены графики решения системы уравнений (30), а также предельные вероятности (31) при $\lambda_1 = 4$, $\mu_1 = 6$, $\lambda_2 = 1$, $\mu_2 = 4$.

В приведенном примере предельные вероятности $P_0 = 0,509$, $P_1 = 0,291$, $P_2 = 0,2$. Как видно на рисунке, графики очень быстро прижимаются к соответствующим асимптотам. Здесь P_0 – доля времени простоя системы, P_1 – доля времени, приходящегося на обслуживание обычных заявок, $P_2 = 1 - P_0 - P_1$ – доля времени, приходящегося на обслуживание приоритетных заявок.

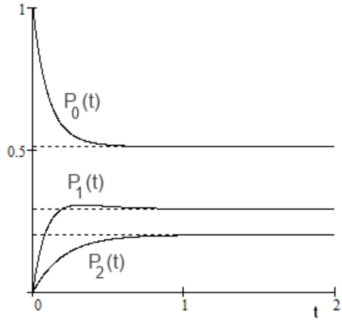

Рис. 9. Вероятности состояний СМО с приоритетным входным потоком

В дальнейшем при исследовании системы мы сконцентрируем внимание на характеристиках, основанных на предельных вероятностях.

Как следует из описания работы СМО, приоритетные заявки ведут себя так, как если бы поток обычных заявок отсутствовал. Таким образом, характеристики обслуживания приоритетных заявок совпадают с характеристиками, рассмотренными в § 2.2. С обычными заявками ситуация несколько иная. Найдем вероятность того, что обслуживание принятой обычной заявки будет завершено до появления приоритетной. Вероятность того, что обычная заявка, находящаяся на обслуживании в момент t, будет обслуживаться в течение элементарного промежутка времени равна dt равна $\mu_1 \cdot e^{-\mu_1 \cdot t} \cdot dt$. Вероятность того, что к моменту t не поступила приоритетная заявка $e^{-\lambda_2 \cdot t}$. По формуле полной вероятности искомая вероятность

$$\int_0^{+\infty} \mu_1 \cdot e^{-(\lambda_2+\mu_1)\cdot t} \cdot t = \frac{\mu_1}{\lambda_2+\mu_1}.$$

Итак, $\mu_1/(\lambda_2+\mu_1)$– вероятность того, что принятая заявка будет обслужена, а, соответственно, $\lambda_2/(\lambda_2+\mu_1)$– вероятность того, что обслуживание принятой заявки будет прервано. Обе вероятности условные, т. е. при условии, что заявка принята. Согласно (31), безусловная вероятность обслуживания обычной заявки равна:

$$P_{\text{обычн\_обсл}} = P_0 \cdot \frac{\mu_1}{(\lambda_2+\mu_1)} = \frac{\mu_1 \cdot \mu_2}{(\lambda_1+\lambda_2+\mu_1)\cdot(\lambda_2+\mu_2)};$$
$$P_{\text{обычн\_прерв}} = P_0 \cdot \frac{\lambda_2}{(\lambda_2+\mu_1)} = \frac{\lambda_2 \cdot \mu_2}{(\lambda_1+\lambda_2+\mu_1)\cdot(\lambda_2+\mu_2)}.$$

Разумеется, $P_{\text{обычн\_обсл}} + P_{\text{обычн\_прерв}} = P_0$ – вероятность принятия обычной заявки на обслуживание.

Характеристики СМО с отказами и приоритетными заявками

1. Ожидаемое **время между двумя последовательными заявками** в обычном потоке

   $\bar{T}_{\text{Ожид\_об}} = \dfrac{1}{\lambda_1}$, в приоритетном потоке $\bar{T}_{\text{Ожид\_пр}} = \dfrac{1}{\lambda_2}$.

2. Ожидаемое **время обслуживания обычной заявки**

$$\bar{T}_{\text{обсл\_об}} = \frac{1}{\mu_1} \text{ и приоритетной } \bar{T}_{\text{обсл\_пр}} = \frac{1}{\mu_2}.$$

3. **Относительная пропускная способность** по приоритетным заявкам $Q_{np} = P_0 + P_1 = \frac{\mu_2}{\lambda_2+\mu_2}$ – доля приоритетных заявок, принимаемых на обслуживание. Все принятые заявки обслуживаются.

4. **Абсолютная пропускная способность** по приоритетным заявкам

$$A_{\text{пр}} = \lambda_2 \cdot Q_{\text{пр}} = \frac{\lambda_2 \cdot \mu_2}{\lambda_2 + \mu_2}$$

– ожидаемое количество обслуживаемых в единицу времени приоритетных заявок.

5. **Относительная пропускная способность** по обычным заявкам

$$Q_{\text{об}} = P_{\text{обыч\_обсл}} = \frac{\mu_2}{\lambda_2 + \mu_2} \cdot \frac{\mu_1}{\lambda_1 + \lambda_2 + \mu_1}$$

– доля обычных заявок, принимаемых на обслуживание. Интересно, что относительная пропускная способность (ОПС) по обычным заявкам равна произведению ОПС по приоритетным заявкам на ту ОПС, которая была бы, если бы отменили приоритеты, т. е. если бы все заявки обслуживались как обыкновенные.

6. **Абсолютная пропускная способность** по обычным заявкам

$$A_{\text{об}} = P_{\text{обыч\_обсл}} \cdot \lambda_1 = \lambda_1 \cdot \frac{\mu_2}{\lambda_2 + \mu_2} \cdot \frac{\mu_1}{\lambda_1 + \lambda_2 + \mu_1}$$

– ожидаемое количество обслуживаемых в единицу времени обычных заявок.

7. **Интенсивность выходящего потока приоритетных заявок, получивших отказ,**

$$P_2 \cdot \lambda_2 = \frac{\lambda_2^2}{\lambda_2 + \mu_2}$$

– ожидаемое количество приоритетных заявок в единицу времени, получающих отказ по причине занятости единственного канала обслуживанием другой приоритетной заявки.

8. **Интенсивность выходящего потока обычных заявок, получивших отказ,**

$$(1 - P_0) \cdot \lambda_1 = (1 - \frac{\mu_2 \cdot (\lambda_2 + \mu_1)}{(\lambda_1 + \lambda_2 + \mu_1)(\lambda_2 + \mu_2)}) \cdot \lambda_1$$

– ожидаемое количество обычных заявок в единицу времени, получивших отказ по причине занятости канала.

9. **Интенсивность выходящего потока обычных заявок, принятых на обслуживание, но не обслуженных по причине появления приоритетной заявки**

$$P_{обыч\_ прер} \cdot \lambda_1 = \frac{\lambda_1 \cdot \lambda_2 \cdot \mu_2}{(\lambda_1 + \lambda_2 + \mu_1)(\lambda_2 + \mu_2)}$$

– количество обычных заявок в единицу времени, вытесненных из системы приоритетными заявками, т. е. заявок, принятых на обслуживание, но необслуженных.

Приведем пример расчета интенсивностей выходящих потоков для представленного на рис. 9 случая: $\lambda_1 = 4$, $\mu_1 = 6$, $\lambda_2 = 1$, $\mu_2 = 4$.

№	Интенсивности выходящих потоков	Обычные заявки	Приоритетные заявки
1	Обслуженных заявок	1,745	0,8
2	Заявок, получивших отказ	1,964	0,2
3	С прерванным обслуживанием	0,291	0
4	ИТОГО	4	1

Очевидно, сумма интенсивностей всех выходящих потоков равна интенсивности соответствующего входящего потока.

Глава 3. Процессы гибели и размножения

Пусть СМО имеет множество состояний $\{S_i\}$, где $i = 0, 1, \ldots, n$. В частности, не исключается случай $n = \infty$. По-прежнему справедливы допущения (18) о вероятностях переходов. При этом возможны только переходы вида

$$\lambda_{i,i+1} \text{ и } \lambda_{i+1,i}, \text{ где } i = 0, 1, \ldots n - 1.$$

Тогда процесс функционирования системы называют **процессом гибели и размножения**. Таким образом, для процессов гибели и размножения характерны только последовательные переходы слева направо или справа налево (рис. 10). Этот класс

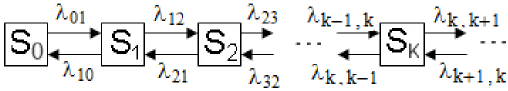

Рис. 10. Процесс гибели и размножения

процессов впервые начали изучать в связи с исследованиями динамики численности популяций, распространения эпидемий и другими подобными задачами. Отсюда и закрепившееся за процессами название. Если возможны переходы только слева направо, процесс называют **процессом чистого размножения**. Если же возможны переходы только в обратном направлении, говорят о **процессе гибели**.

Запишем уравнения Колмогорова для произвольной системы гибели и размножения. Согласно схеме первое уравнение

$$P_0'(t) = -\lambda_{01} \cdot P_0(t) + \lambda_{10} \cdot P_1(t),$$

далее

$$P_k'(t) = \lambda_{k-1,k} \cdot P_{k-1}(t) - (\lambda_{k,k-1} + \lambda_{k,k+1}) \cdot P_k(t) + \lambda_{k+1,k} \cdot P_{k+1}(t),$$

где $k > 0$. Если множество состояний конечно и номер крайнего справа – n, то систему будет замыкать уравнение

$$P_n'(t) = \lambda_{n-1,n} \cdot P_{n-1}(t) - \lambda_{n,n-1} \cdot P_n(t).$$

Уравнения Колмогорова для случая процессов гибели и размножения иногда называют уравнениями Эрланга.

§ 3.1. Формулы Эрланга

Аналитическое решение систем уравнений Колмогорова даже для простых процессов гибели и размножения с конечным числом состояний часто оказывается довольно громоздким. Однако если существует установившееся решение, получить его нетрудно. Система уравнений Колмогорова для установившегося решения принимает вид

$$\begin{cases} -\lambda_{01} \cdot P_0 + \lambda_{10} \cdot P_1 = 0; \\ \lambda_{01} \cdot P_0 - (\lambda_{10} + \lambda_{12}) \cdot P_1 + \lambda_{21} P_2 = 0; \\ \dots \\ \lambda_{k-1,k} \cdot P_{k-1} - (\lambda_{k,k-1} + \lambda_{k,k+1}) \cdot P_k + \lambda_{k+1,k} P_{k+1} = 0. \end{cases}$$

Положим, $z_k = \lambda_{k,k+1} \cdot P_k - \lambda_{k+1,k} \cdot P_{k+1}$, тогда

$$z_0 = 0, \qquad z_0 - z_1 = 0, \qquad \dots \qquad z_{k-1} - zk = 0 \qquad .$$

Отсюда, независимо от того, конечна система или нет,

$$z_0 = z_1 = \dots = zk = \dots = 0$$

и, следовательно, $P_{k+1} = \frac{\lambda_{k,k+1}}{\lambda_{k+1,k}} \cdot P_k$, где $k > 0$. По индукции получим выражение предельных вероятностей через

$$P_k = (\prod_{i=0}^{k-1} \lambda_{i,i+1} / \prod_{i=0}^{k-1} \lambda_{i+1,i}),$$

где $k > 0$. Введем обозначение:

$$\alpha_k = (\prod_{i=0}^{k-1} \lambda_{i,i+1} / \prod_{i=0}^{k-1} \lambda_{i+1,i}),$$

где $k > 0$. Тогда $P_k = \alpha_k \cdot P_0$. Таким образом, чтобы получить α_k, надо просто произведение интенсивностей всех переходов, ведущих на схеме 10 слева направо к S_k, разделить на произведение интенсивностей всех переходов, ведущих справа налево от S_k . Если положить $\alpha_0 = 0$, то из условия $\sum_{k=0}^{n} P_k = 1$ непосредственно следует

$$\sum_{k=0}^{n} \alpha_k \cdot P_0 = 1 \qquad \Longrightarrow \qquad P_0 = (\sum_{k=0}^{n} \alpha_k)^{-1}.$$

Все сказанное справедливо и для случая $n = \infty$, если соответствующая сумма сходится. Итак,

$$\alpha_k = \begin{cases} 1, \ \textit{если} \ k = 0; \\ (\prod_{i=0}^{k-1} \lambda_{i,i+1} / \prod_{i=0}^{k-1} \lambda_{i+1,i}), \ \textit{если} \ k > 0. \end{cases} \tag{32}$$

$$P_k = \alpha_k \cdot (\sum_{k=0}^{n} \alpha_k)^{-1}. \tag{33}$$

Выражения (32–33) известны как формулы Эрланга, поскольку именно Эрланг впервые получил их для установившегося процесса в многоканальной СМО с отказами. Ниже мы рассмотрим примеры использования формул (32 - 33) для различных типов СМО.

§ 3.2. Многоканальная СМО с отказами

Система с отказами и n каналами обслуживания имеет конечное множество состояний $\{S_i\}$, где $i = 0, 1, \ldots, n$: S_0 – свободны все каналы, S_1 – занят один канал, S_2 – заняты два канала и так далее, S_n – заняты все n каналов обслуживания. Интенсивность входящего потока – λ , интенсивность обслуживания – μ. Как видно на схеме (рис. 11), интенсивности переходов по возрастанию индекса совпадают с интенсивностью входящего потока, а интенсивности переходов по убыванию индекса зависят от индекса состояния.

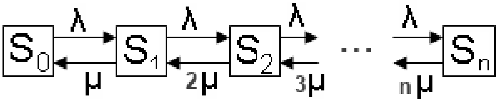

Рис. 11. Многоканальная СМО с отказами

Так, интенсивность перехода $S_k \to S_{k-1}$ равна $k \cdot \mu$, т. е. произведению интенсивности обслуживания одним каналом μ на количество задействованных каналов k.

Применим формулы Эрланга $\alpha_0 = 1$ и $\alpha_k = \frac{\lambda_k}{1 \cdot 2 \cdot \ldots \cdot k \cdot \mu^k}$ для k>0. Введем обозначения $\rho = \frac{\lambda}{\mu}$ – приведенная интенсивность входящего потока или нагрузка системы.
Тогда

$$\alpha_k = \frac{\rho^k}{k!},\ где\ k = 0, 1, \ldots, n, \qquad P_0 = (\sum_{k=0}^{n} \frac{\rho^k}{k!})^{-1}.$$

Отказ в обслуживании заявка получает только тогда, когда система находится в состоянии S_n, т. е. все каналы заняты. Вероятность этого события P_n.

Характеристики многоканальной СМО с отказами

1. **Относительная пропускная способность системы**.

$$Q = 1 - P_n$$

2. **Абсолютная пропускная способность, она же интенсивность выходящего потока обслуженных заявок**

$$A = Q \cdot \lambda = (1 - P_n) \cdot \lambda.$$

3. **Интенсивность выходящего потока заявок, получивших отказ,**

$$P_n \cdot \lambda.$$

4. **Среднее число занятых обслуживанием каналов, т. е. математическое ожидание числа занятых каналов**

$$\begin{aligned}
\bar{K} &= M(k) = \sum_{k=0}^{n} k \cdot P_k = \sum_{k=1}^{n} k \cdot P_k = \sum_{k=1}^{n} k \frac{\rho^k}{k!} \cdot P_0 = \\
&= \rho \cdot \sum_{k=1}^{n} \frac{\rho^{k-1}}{(k-1)!} \cdot P_0 = \rho \cdot \sum_{k=0}^{n-1} \frac{\rho^k}{k!} \cdot P_0 = \\
&= \rho \cdot \sum_{k=0}^{n-1} P_k = \rho \cdot (1 - P_n).
\end{aligned}$$

Таким образом, среднее число занятых каналов равно произведению приведенной интенсивности входящего потока на относительную пропускную способность системы.

5. **Среднее число свободных от обслуживания каналов**

$$\bar{N_0} = n - \bar{K}.$$

6. **Коэффициент занятости каналов** $\bar{K}/n$.

7. **Коэффициент простоя каналов** $\bar{N_0}/n$.

Иногда рассматривают СМО с бесконечным числом каналов. Хотя кто-то может вполне резонно возразить, что таковых не бывает, названная модель вполне адекватно описывает некоторые реальные ситуации. Например, если каналы обслуживания – номера в гостинице на курорте в мертвый сезон, а также в других ситуациях, когда каналов очень много, а заявок очень мало. Очевидно, в такой системе обслуживаются все заявки.

В этом случае

$$P_0 = (\sum_{k=0}^{\infty} \frac{\rho^k}{k!})^{-1} = e^{-\rho}, \qquad P_k = \frac{\rho^k}{k!} \cdot e^{-\rho}.$$

Среднее число занятых каналов

$$\begin{aligned} \bar{K} &= \sum_{k=0}^{\infty} k \cdot P_k = \sum_{k=1}^{\infty} k \frac{\rho^k}{k!} \cdot e^{-\rho} = \rho \sum_{k=1}^{\infty} \frac{\rho^{k-1}}{(k-1)!} \cdot e^{-\rho} = \\ &= \rho \cdot \sum_{k=1}^{\infty} (\frac{\rho^k}{k!})'_{\rho} \cdot e^{-\rho} = \rho \cdot (\sum_{k=1}^{\infty} \frac{\rho^k}{k!})'_{\rho} \cdot e^{-\rho} = \rho \cdot (e^{\rho})'_{\rho} \cdot e^{-\rho} = \rho. \end{aligned}$$

Таким образом, среднее число занятых каналов в «бесконечно-канальной» СМО $\bar{K} = \rho$ определяется только отношением интенсивности входящего потока к интенсивности обслуживания.

§ 3.3. Одноканальная СМО без ограничений на длину очереди

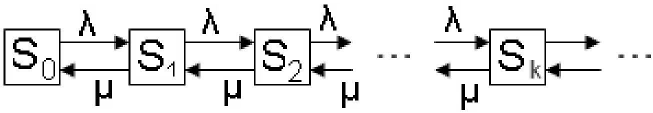

Рис. 12. Одноканальная СМО без ограничений на длину очереди

Одноканальная система без ограничений на длину очереди (рис. 12) имеет бесконечное множество состояний $\{S_i\}$, где где $i = 0, 1, \dots :$ S_0 – единственный канал свободен, S_1 – канал занят обслуживанием заявки, S_2 – канал занят, и одна заявка находится в очереди, S_3 – канал занят, и две заявки находятся в очереди и т. д. Таким образом, состояние S_k , где $k > 1$, – это когда в очереди находится $k - 1$ заявка. Интенсивность входящего потока – λ, интенсивность обслуживания – μ. Интенсивности всех переходов в порядке возрастания индекса равны λ, а интенсивности всех переходов в обратном направлении, в отличие от случая из § 3.2, совпадают и равны μ .

Если заканчивается обслуживание очередной заявки при наличии очереди, система переходит к обслуживанию следующей, а длина очереди уменьшается на единицу.

Формулы Эрланга теперь примут вид

$$\alpha_k = \frac{\lambda^k}{\mu^k} = \rho^k, \text{ где } k = 0, 1, 2 \ldots; \qquad P_0 = (\sum_{k=0}^{\infty} \rho^k)^{-1} = 1 - \rho.$$

Разумеется, должно выполняться условие $\rho < 1$. В противном случае, когда $\lambda \geq \mu$, интенсивность обслуживания не превышает интенсивности поступления заявок, очередь со временем стремится к бесконечности и сам разговор о предельных вероятностях теряет смысл.

В дальнейшем мы не будем специально выделять основные характеристики рассматриваемых СМО.
Вероятности состояний системы $P_k = (1 - \rho) \cdot \rho^k$. Любая поступившая заявка принимается в очередь и рано или поздно обслуживается. Отсюда относительная пропускная способность $Q = 1$, а абсолютная $A = \lambda$. Выходящий поток только один и образован обслуженными заявками, его интенсивность равна интенсивности входящего потока λ.

Найдем среднюю длину очереди $\bar{L}_{оч}$ для установившегося решения. Вероятность нулевой длины равна $P_0 + P_1$, и далее при $k > 1$

вероятность длины k равна P_{k+1}. Таким образом,

$$\begin{aligned}
\bar{L}_{оч} &= M(k) = \sum_{k=1}^{\infty} k \cdot P_{k+1} = \\
&= (1-\rho)\sum_{k=1}^{\infty} k \cdot \rho^{k+1} = (1-\rho)\cdot\rho^2\cdot\sum_{k=1}^{\infty} k\cdot\rho^{k-1} = \\
&= (1-\rho)\cdot\rho^2\cdot(\sum_{k=1}^{\infty}\rho^k)'_\rho = (1-\rho)\cdot\rho^2\cdot(\frac{\rho}{1-\rho})'_\rho = \frac{\rho^2}{1-\rho}.
\end{aligned}$$

Пусть $T_{оч}$ – время нахождения вновь поступившей заявки в очереди. $M(T_{оч}/S_k)$ – ожидаемое время пребывания заявки в очереди, при условии, что система на момент поступления заявки находится в состоянии S_k. Тогда

$$M(T_{оч}/S_k) = \begin{cases} 0, & \textit{если } k = 0; \\ \frac{k}{\mu}, & \textit{если } k > 0. \end{cases}$$

По формуле полной вероятности ожидаемое время пребывания заявки в очереди

$$\begin{aligned}
\bar{T}_{оч} &= \sum_{k=0}^{\infty} P(S_k)\cdot M(T_{оч}/S_k) = \sum_{k=1}^{\infty} P_k\cdot M(T_{оч}/S_k) = \\
&= \sum_{k=1}^{\infty} P_k\cdot\frac{k}{\mu} = \frac{1-\rho}{\mu}\cdot\sum_{k=1}^{\infty} k\cdot\rho^k = \frac{\rho}{\mu(1-\rho)}.
\end{aligned}$$

Отсюда – первая формула Литтла: $\bar{L}_{оч} = \lambda\cdot\bar{T}_{оч}$.

Среднее время пребывания заявки в системе равно сумме среднего времени пребывания в очереди и среднего времени обслу-

живания:
$$\bar{T}_{сис} = \bar{T}_{оч} + \bar{T}_{обс} = \frac{\rho}{\mu(1-\rho)} + \frac{1}{\mu}.$$
Среднее число заявок, находящихся под обслуживанием, $\bar{L}_{обсл} = 1 - P_0 = \rho$. Тогда среднее число находящихся в системе заявок

$$\bar{L}_{сис} = \bar{L}_{оч} + \bar{L}_{обс} = \frac{\rho^2}{1-\rho} + \rho = \frac{\rho}{1-\rho}.$$

Отсюда легко получается соотношение между количеством заявок в системе и временем пребывания заявки в системе (34) – вторая формула Литтла. Таким образом,

$$\bar{L}_{оч} = \lambda \cdot \bar{T}_{оч}, \qquad \bar{L}_{сис} = \lambda \cdot \bar{T}_{сис}. \tag{34}$$

Рассмотренная нами в этом параграфе система без ограничений на длину очереди является довольно распространенной моделью реальных систем. Например, пусть пропускная способность городского травматологического пункта – 10 пациентов в час, а в городе в этот период времени случается 9 травм в час. Чему равна средняя длина очереди $\bar{L}_{оч}$ и среднее время пребывания пациента в очереди $\bar{T}_{оч}$? Итак: $\lambda = 9$, $\mu = 10$. Следовательно, $\rho = 0,9$; $\bar{L}_{оч} = 8,1$ пациента; $\bar{T}_{оч} = 0,9$ часа, или 54 минуты. При этом, если смена длится 6 часов, то из них 0,6 часа, или 36 минут, пункт простаивает. Эти моменты важно учесть при подготовке организационных решений, связанных с медицинским обслуживанием и другими видами обслуживания населения.

§ 3.4. Одноканальная СМО с ограничением на длину очереди

Теперь рассмотрим систему, аналогичную исследованной в предыдущем параграфе, но с ограничением на длину очереди. Пусть система имеет $n+1$ состояние: S_0 – единственный канал свободен, S_1 – единственный канал занят обслуживанием заявки, S_2 – канал занят и одна заявка находится в очереди и т. д. Последнее состояние S_{n+1} – в очереди n заявок (рис. 13).

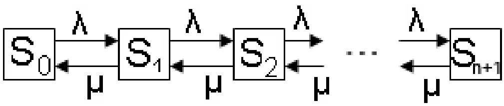

Рис. 13. Одноканальная СМО с ограничением на длину очереди

Формулы Эрланга:

$$\alpha_k = \frac{\lambda^k}{\mu^k} = \rho^k, \text{ } где \text{ } k = 0, 1, 2 \ldots, n+1; \qquad P_0 = (\sum_{k=0}^{n+1} \rho^k)^{-1}.$$

Выполнение условия $\rho < 1$ теперь не требуется. $P_k = P_0 \cdot \rho^k$. Вероятность принятия заявки на обслуживание равна $1 - P_{n+1}$, вероятность отказа – P_{n+1}. Система имеет два выходных потока: поток обслуженных заявок с интенсивностью $A = (1 - P_{n+1})\lambda$ и поток заявок, получивших отказ, с интенсивностью $A = P_{n+1} \cdot \lambda$. Средняя длина очереди

$$\bar{L}_{оч} = M(k) = \sum_{k=1}^{n} k \cdot P_{k+1} = P_0 \cdot \sum_{k=1}^{n} k \cdot \rho^{k+1}. \qquad (35)$$

Аналогично тому, как это делалось в предыдущем параграфе, нетрудно найти $\bar{L}_{обс}$, $\bar{L}_{сис}$, $\bar{T}_{оч}$, $\bar{T}_{сис}$. Разумеется, формулы Литтла также выполняются.

§ 3.5. Одноканальная СМО с нетерпеливыми заявками

Описание системы совпадает с представленным в § 3.2. Единственное отличие – новый выходящий поток, поток нетерпеливых заявок. Для него мы вводим новый параметр ω – интенсивность ухода заявки из очереди. Таким образом, $\frac{1}{\omega}$ – среднее время ожидания заявки в очереди. Схема СМО изображена на рис. 14.

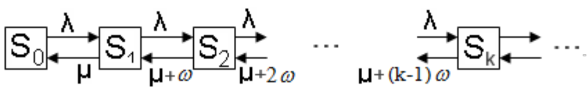

Рис. 14. Одноканальная СМО с нетерпеливыми заявками

В рассматриваемом примере заявки иногда покидают очередь по своей инициативе, не дождавшись обслуживания. Формулы Эрланга:

$$\begin{aligned} \alpha_0 &= 1, \alpha_k = \frac{\lambda^k}{\prod_{i=1}^{k}(\mu + (i-1)\cdot\omega)} = \rho^k,\ \textit{для}\ k > 0; \\ P_0 &= (\sum_{k=0}^{n+1} \rho^k)^{-1}, P_k = \alpha_k \cdot P_0. \end{aligned}$$

Также рассматривают приведенную интенсивность потока уходов $\beta = \frac{\omega}{\mu}$. Тогда

$$\alpha_k = \frac{\rho^k}{\prod_{i=1}^{k}(1 + (i-1) \cdot \beta)}.$$

Средняя длина очереди $\bar{L}_{оч} = M(k) = \sum_{k=1}^{n} k \cdot P_{k+1}$. Каждая заявка «испытывает желание» уйти с интенсивностью ω . Поэтому интенсивность выходящего потока покинувших очередь заявок равна $\bar{L}_{оч} \cdot \omega$. Поскольку в очередь принимаются все без исключения заявки, абсолютная пропускная способность системы $A = \lambda - \bar{L}_{оч} \cdot \omega$.

§ 3.6. Замкнутая одноканальная СМО

До сих пор мы рассматривали СМО, в которых входной поток не зависит от того, сколько заявок в данный момент обслуживается или находится в очереди. Так, в большом городе вы можете считать, что ваш звонок не окажет влияния на общую интенсивность звонков по городу и, хотя население даже очень большого города конечно, вы можете в модели считать количество источников заявок бесконечным. Другое дело, например, когда в цехе всего десять станков и один мастер по их ремонту. Тогда моделью цеха будет СМО с состояниями $\{S_i\}$, где $i = 0, 1, \ldots, n : S_0$ – работают все станки, S_1 – один станок в ремонте, остальные работают, S_2 – один в ремонте, один в очереди, остальные работают и, наконец, S_n– один в ремонте, остальные в очереди на ремонт! Интенсивность потока заявок на ремонт с одного станка, т. е. интенсивность потока отказов – λ , интенсивность обслуживания –

μ . Величину $1/\lambda$ в теории надежности называют наработкой на отказ. Такие системы называют замкнутыми, или системами Энгсета (рис. 15).

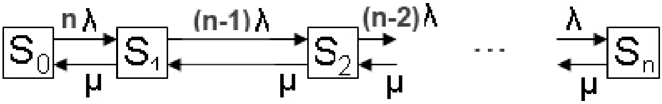

Рис. 15. Замкнутая одноканальная СМО

Формулы Эрланга:

$$\begin{aligned} \alpha_0 &= 1 \text{ и } \alpha_k = \frac{k!}{(n-k)!} \cdot \rho^k, \text{ для } k = 1, 2 \ldots n; \\ P_0 &= (\sum_{k=0}^{n} \rho^k)^{-1}, P_k = \alpha_k \cdot P_0. \end{aligned}$$

Абсолютная пропускная способность СМО $A = P_{зан} \cdot \mu$, где $P_{зан}$ – вероятность того, что система занята обслуживанием заявки. Напомним, что здесь μ – производительность системы при условии, что она обслуживала бы заявки непрерывно, без простоев. В замкнутой СМО можно выделить некоторые количества активных $N_{акт}$ и пассивных $N_{пас}$ источников заявок. Например, пассивные источники – это находящиеся в ремонте или в очереди на ремонт станки. Только активный источник может подать новую заявку на обслуживание. Очевидно, $N_{акт} + N_{пас} = n$. Также и для средних: $\bar{N}_{акт} + \bar{N}_{пас} = n$. Средняя интенсивность входящего потока

$$\bar{\Lambda} = A = (1 - P_0) \cdot \mu = (n - \bar{N}_{пас}) \cdot \lambda \implies \qquad \bar{N}_{пас} = n - \frac{1 - P_0}{\rho}.$$

Здесь

$$\bar{N}_{\,nac} \quad = \quad L_{cuc};$$

$$\begin{aligned}\bar{L}_{o\text{ч}} \quad &= \quad \bar{L}_{cuc} - \bar{L}_{o\text{б}c\text{л}} = \\ &= \quad n - \frac{1-P_0}{\rho} - (1-P_0) = n - (1-P_0)\cdot(\frac{1}{\rho}+1).\end{aligned}$$

Вероятность того, что заявка активна, т. е. доля активных заявок

$$P_{a\kappa m} = 1 - \frac{\bar{L}_{cuc}}{n\cdot\lambda}; \qquad \bar{T}_{o\text{ч}} = \frac{\bar{L}_{o\text{ч}}}{\bar{\Lambda}}.$$

* * *

К сожалению, объем пособия не позволил рассмотреть многоканальные системы с ограничением и без ограничений на длину очереди, многоканальные системы с нетерпеливыми заявками и замкнутые многоканальные системы. Также не рассмотрены многофазовые системы, пропущен ряд теоретических вопросов существования решений уравнений Колмогорова, а также особые случаи входящего потока и порядка обслуживания, случаи, когда нарушается ординарность потока или когда заявки поступают по законам простейшего потока, но отдельные заявки не обслуживаются. Например, занятия на курсах английского языка начинаются по мере комплектования групп. В таких случаях не всегда классические учебники дадут ответ на все вопросы и приходится обращаться к дополнительной литературе. Однако студент, освоивший методы моделирования СМО и основные характеристики СМО на рассмотренных в книге примерах и вооружившись необходимой литературой, сможет самостоятельно исследовать многие пропущенные нами классы СМО. В конце концов для чего-то существуют и «толстые» книги.

Биографические справки

1. Вентцель Елена Сергеевна (1907–2002) – советский математик, доктор технических наук, профессор, автор научных работ и учебников по теории вероятностей, исследованию операций и теории массового обслуживания. Елена Сергеевна является также автором ряда романов, повестей и рассказов.

2. Гнеденко Борис Владимирович (1912–1995) – советский математик, академик АН УССР, специалист по теории вероятностей и математической статистике. Занимался задачами, связанными с обороной страны. Автор многих работ по теории массового обслуживания, один из основоположников теории надежности.

3. Йохансон Фредерик Фердинанд (1855–1934) – датский инженер, директор Копенгагенской телефонной компании. Изложил первые идеи теории очередей в 1907 году в статье «Время ожидания и число вызовов».

4. Каштанов Виктор Алексеевич (1934) – советский математик, доктор физико-математических наук, лауреат Государственной премии СССР 1979 года. Один из руководителей разработок в области фундаментальных проблем системной безопасности.

5. Коваленко Игорь Николаевич (1935) – советский математик, доктор технических наук, доктор физико - математических наук, профессор, академик АН УССР, лауреат Государственной премии СССР 1978 года. Основные труды относятся к теории вероятностей и ее приложениям, теории надежности, теории массового обслуживания.

6. Колмогоров Андрей Николаевич (1903–1987) – выдающийся советский математик, академик АН СССР, Герой Социалистиче-

ского Труда, лауреат Ленинской и Государственной премий СССР, профессор МГУ, иностранный член Национальной академии наук США, член Лондонского королевского общества, член Германской академии естествоиспытателей и многих других академий и научных обществ мира. Внес значительный вклад в развитие теории марковских процессов с непрерывным временем.

7. Клейнрок Леонард (1934) – американский инженер в области информационных технологий и компьютерных сетей. Сыграл важную роль в создании сети ARPANET – предшественницы Интернета.

8. Марков Андрей Андреевич (1856–1922) – русский математик, академик Российской академии наук, профессор физико-математического факультета Санкт- Петербургского университета. Внес большой вклад в теорию вероятностей, математический анализ и теорию чисел, впервые исследовал широкий класс стохастических процессов с дискретным временем (марковские цепи) и непрерывным временем (марковские процессы). Цепи Маркова нашли широкое применение в работах Планка, Эйнштейна и других ученых.

9. Пуассон Симеон Денни (1781–1840) – французский математик, механик и физик, профессор Политехнической школы в Париже, почетный член Петербургской академии наук. Предложил один из важнейших законов распределения случайных величин, впоследствии названный его именем.

10. Хинчин Александр Яковлевич (1894–1959) – советский математик, один из наиболее значимых ученых в советской школе теории вероятностей, член-корреспондент АН СССР, действительный член и один из основателей Академии педагогических наук,

лауреат Сталинской премии 1941 года. Является (совместно с А. Н. Колмогоровым) создателем современной теории случайных процессов и теории массового обслуживания. В молодости Александр Яковлевич опубликовал четыре небольших сборника своих стихов.

11. Энгсет Торе Олаус (1865–1943) – норвежский математик и инженер, один из основоположников теории массового обслуживания.

12. Эрланг Агнер Краруп (1878–1929) – датский математик, статистик и инженер, сотрудник Копенгагенской телефонной компании, основатель научного направления по изучению трафика в телекоммуникационных системах. В 1909 году Эрланг опубликовал свою первую работу «Теория вероятностей и телефонные разговоры», которая была признана во всем мире и легла в основу теории массового обслуживания.

Список литературы

1. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989. 540с.

2. Вентцель Е. С., Овчаров Л. А. Теория случайных процессов и ее инженерные приложения. М.: ACADEMA, 2003. 432 с.

3. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. М.: Наука, 1987. 336 с.

4. Ивченко Г. И., Каштанов В. А., Коваленко И. Н. Теория массового обслуживания. М.: Высшая школа, 1982. 256 с.

5. Клейнрок Л. Теория массового обслуживания. М.: Машиностроение, 1979. 432 с.

6. Новиков О. А., Петухов С. И. Прикладные вопросы теории массового обслуживания. М.: Советское радио, 1969. 400 с.

7. Лабскер Л. Г., Бабешко Л. О. Теория массового обслуживания в экономической сфере. М.: Издательское объединение ЮНИТИ, 1999. 319 с.

8. Хинчин А. Я. Работы по математической теории массового обслуживания. М.: Государственное издательство физико-математической литературы, 1963. 236 с.

9. Чернецкий В. И. Математическое моделирование стохастических систем. Петрозаводск: Издательство Петрозаводского государственного университета, 1994. 488 с.

Учебное издание

Белый Евгений Константинович

Введение в теорию массового обслуживания

Учебное пособие для студентов, обучающихся по направлению «Информационные системы и технологии»

Редактор *Е. Е. Порывакина*
Компьютерная верстка *Е. К. Белого*

Фотография для обложки предоставлена старшим преподавателем
кафедры физической культуры
Г. А. Крикуновым

Подписано в печать 17.12.14. Формат 60x84 1/16.
Бумага офсетная. Уч.-изд. л. 2,7. Тираж 105 экз. Изд. № 168

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
ПЕТРОЗАВОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Отпечатано в типографии Издательства ПетрГУ
185910, Петрозаводск, пр. Ленина, 33

ISBN: 978-5-8021-2203-7

9 785802 122037