Министерство образования РФ

Московский Государственный Технический Университет

им. Н. Э. Баумана

Кафедра ИУ3 «Информационные системы и телекоммуникации»

ОСНОВЫ ТЕОРИИ НАДЕЖНОСТИ

Конспект лекций

Составил доц. ктн Герман ДЯ

2005 г.

1. Введение

Настоящая работа представляет собой первую редакцию конспекта лекций по курсу «Надежность информационных систем».

Курс разделен на четыре раздела. В первом разделе приводятся основные понятия теории надежности. Рассматриваются основные законы надежности, причем основное внимание уделяется экспоненциальному закону. Рассматривается надежность систем с последовательным (основным) и параллельным соединением элементов. Указывается на роль резервирования как важного средства повышения надежности систем, перечисляются различные способы резервирования.

Во втором разделе вводится аппарат уравнений Колмогорова для марковских систем с дискретными состояниями и непрерывным временем. Этот аппарат очень прост и удобен для решения различных инженерных задач. Здесь рассмотрены некоторые задачи определения надежности резервированных систем, систем с восстановлением отказавших элементов. В этом же разделе рассматриваются основные понятия систем массового обслуживания (СМО) и рассмотрены некоторые вводные задачи для них постольку, поскольку для их решения используются те же уравнения Колмогорова. Рассмотрены два подхода к расчету буферных запоминающих устройств (БЗУ).

В третьем разделе рассматриваются в рамках аппарата булевой алгебры некоторые способы построения контрольных и диагностических тестов.

В четвертом разделе рассматриваются вопросы тестирования программных изделий. В этом разделе содержаться материалы, использование которых обязательно при выполнении дипломного проекта.

В дальнейшем по мере появления в литературе конкретных материалов по определению характеристик надежности программных продуктов они должны быть обязательно включены в настоящий конспект. Вопросы тестирования программных продуктов рассматриваются в других курсах специальности.

2. Основные понятия и определения теории надежности

Под надежностью понимается свойство изделия (детали, узла, системы) выполнять свои функции при определенных условиях эксплуатации в течение требуемого периода времени. Теория надежности изучает процессы возникновения отказов элементов и систем, создает математические модели, методы их экспериментального подтверждения, находит пути и средства повышения надежности.

Все объекты в теории надежности разбиваются на системы и элементы, из которых состоят системы. Вначале будем изучать надежность элементов, при этом предполагаем, что состояние изучаемого элемента не зависит от состояния остальных элементов. Рассматриваются элементы непрерывного действия.

Под отказом понимаем переход элемента от исправного состояния к неисправному состоянию. Предполагаем, что этот переход происходит мгновенно, в какой-то момент времени τ. Очевидно, что τ, которое называется временем жизни элемента, есть непрерывная случайная величина.

Под *вероятностью отказа* примем вероятность того, что время жизни меньше, чем рассматриваемый интервал времени

$$\mathbf{Q(t) = P(\tau < t)} \tag{2.1}$$

То есть отказ произойдет на интервале времени $\mathbf{0...t}$. Выражение (2.1) есть интегральный закон распределения τ.

Вероятность безотказной работы называется функцией надежности и равна

$$\mathbf{P(t) = 1 - Q(t) = P(\tau > t)} \tag{2.2}$$

Это вероятность того, что отказ произойдет вне рассматриваемого интервала времени, до момента времени $\mathbf{t}$ отказа не будет. Рассмотрим некоторые свойства функции надежности:

1. $\mathbf{P(0)=1}$. В момент своего первого включения элемент исправен Известно, что это предположение не всегда выполняется, но предположим, что выходной контроль действует исправно и бракованных элементов в системе нет.

2. $\mathbf{P(\infty)=0}$. Нет «вечных» элементов.

3. $\mathbf{P(t)}$ есть непрерывная функция времени. Нет «роковых» мгновений, точно в которые происходят удары, обрывы, разрушения.

4. Функция надежности монотонно убывает, так как $\mathbf{Q(t)=1-P(t)}$ интегральный закон распределения непрерывной случайной величины монотонно возрастает.

Если $\mathbf{Q(t)}$ и $\mathbf{P(t)}$ непрерывные функции, то существует *плотность вероятности отказов*

$$\mathbf{q(t)=\frac{dQ(t)}{dt}=-\frac{dP(t)}{dt}} \tag{2.3}$$

которая имеет второе название – *частота отказов*.

Для пояснения этого названия представим себе опыт по экспериментальной статистической оценке частоты отказов.

Включаются одновременно $\mathbf{N_0}$ однородных элементов.

Фиксируется число элементов $\mathbf{\Delta n_1}$, отказавших на интервале времени $\mathbf{0...\Delta t}$

Фиксируется число элементов $\mathbf{\Delta n_2}$, отказавших на интервале времени $\mathbf{\Delta t...2\Delta t}$

Тогда $\mathbf{\frac{\Delta n_1}{N_0}}$ есть оценка вероятности отказа на первом интервале, а величина $\mathbf{\frac{\Delta n_1}{N_0 \Delta t}}$ есть оценка плотности вероятности отказов или частоты отказов. Величина $\mathbf{\frac{\Delta n_1}{\Delta t}}$ это частота отказов (отношение числа событий к интервалу времени). То есть плотность вероятности отказов это отношение частоты отказов к числу включенных элементов.

Оценкой плотности вероятности будет величины

$$\hat{q} = \frac{\Delta n_i}{N_0 \Delta t} \tag{2.4}$$

Если предыдущие характеристики относились ко всем элементам, то следующая характеристика будет учитывать только те элементы, которые остались исправны в рассматриваемый момент времени. Определим вероятность того, что элемент, исправный в момент времени t_1, останется исправным к моменту времени t_2. Обозначим эту вероятность $Q(t_2 \mid t_1)$. Вертикальная черта в списке аргументов обозначает как обычно, что вероятность условная. Условием является – в момент времени t_1 элемент исправен.

Можно записать, что вероятность безотказной работы на интервале $0...t_2$ есть произведение вероятностей совместного появления двух событий: элемент проработал без отказа до момента времени t_1 и второго события – элемент проработал без отказа на интервале $0...t_2$

$$P(t_2) = P(t_1)P(t_2 \mid t_1) \tag{2.5}$$

Следовательно, условная вероятность безотказной работы на интервале $t_1...t_2$ равна

$$P(t_1 \mid t_2) = \frac{P(t_2)}{P(t_1)} \tag{2.6}$$

Условная вероятность отказа элемента на интервале $t_1...t_2$ равна

$$Q(t_2 \mid t_1) = 1 - P(t_2 \mid t_1) = 1 - \frac{P(t_2)}{P(t_1)} = \frac{P(t_1) - P(t_2)}{P(t_1)}$$

Если интервал времени $\Delta t = t_2 - t_1$ мал, то

$$Q(t_2 \mid t_1) \approx -\frac{P'(t)}{P(t)}\Big|_{t=t_1} \Delta t \tag{2.7}$$

Введем обозначение

$$\lambda(t) = -\frac{P'(t)}{P(t)} \tag{2.8}$$

Тогда

$$Q(t + \Delta t \mid t) = \lambda(t)\Delta t \tag{2.9}$$

Функция $\lambda(t)$, если она существует, называется *опасностью отказа* и играет основную роль при расчете надежности. Она определяет степень надежности элемента в каждый момент времени, но в отличие от частоты отказов, которая является априорной плотностью вероятности отказов, опасность отказов является условной плотностью вероятности отказов. Условие заключается в том, что в процессе вывода этой характеристики мы предполагали, что элемент в текущий момент времени исправен, не успел отказать. То есть,

$q(t)\Delta t$ есть вероятность отказа элемента на малом интервале Δt , вычисленная до того как элемент был включен в работу, или при отсутствии информации о его состоянии в процессе работы (вероятность смерти людей от 70 до 71 года).

$\lambda(t)\Delta t$ есть вероятность отказа элемента на малом интервале Δt, вычисленная для тех элементов, которые до этого момента не отказали (вероятность смерти людей, доживших до 70 лет, в течении ближайшего года).

Опасность отказа более точно описывает процесс эксплуатации с точки зрения надежности, чем частота отказа.

Пусть было включено 1000 элементов. Через час вышло из строя 50 элементов. После 30 часов работы осталось 100 элементов и за ближайший час отказало 20 элементов.

Если судить по априорной плотности вероятности – частоте отказа, то она в процессе эксплуатации уменьшилась. В начале эксплуатации она равна

50/1000, а в конце 20/1000. Однако элемент в процессе эксплуатации стал хуже. В начале опасность отказа была равна 50/1000 (отказывало 5%) , а в конце опасность отказа стала равна 20/100 (отказывало 20%).

Заметим, что если опасность отказа $\lambda(t)$ существует, то функция надежности или закон надежности определяется согласно (2.8), как решение дифференциального уравнения

$$P'(t) = -\lambda(t)P(t),$$

которое можно представить в виде

$$P(t) = e^{-\int \lambda(t)dt} \tag{2.10}$$

Определим еще некоторые характеристики теории надежности. *Среднее время безотказной работы* (средняя наработка на отказ) есть математическое ожидание случайной величины τ - времени отказа. Имеем

$$T_0 = M[\tau] = \int_0^\infty tq(t)dt = -tP(t)|_0^\infty + \int_0^\infty P(t)dt = \int_0^\infty P(t)dt \tag{2.11}$$

Первое слагаемое при вычислении интеграла по частям равно нулю, так как все принятые законы надежности $P(t)$ убывают на бесконечности быстрее чем $\frac{1}{t^2}$.

Дисперсия времени безотказной работы равна

$$D_\tau = M[(\tau - T_0)^2] = M[\tau] - T_0{}^2 = \int_0^\infty t^2 q(t)dt - T_0{}^2 = -t^2 P(t)|_0^\infty +$$

$$+ 2\int_0^\infty tP(t)dt - T_0{}^2$$

Первое слагаемое в окончательном результате равно нулю по причинам, указанным при выводе (2.11). Тогда

$$D_\tau = 2\int_0^\infty tP(t)dt - T_0{}^2 \tag{2.12}$$

Контрольные вопросы

1. Что такое вероятность отказа, вероятность безотказной работы, плотность вероятности отказа?

2. В чем разница между частотой отказа и опасностью отказа?

3. Что такое среднее время безотказной работы?

3. Экспоненциальный закон надежности

Зависимость опасности отказов от времени очевидно может быть трех типов:

1. Опасность отказов с увеличением времени эксплуатации убывает. Элемент в процессе эксплуатации улучшает свои надежностные характеристики. Этот редкий случай имеет место, например, в процессе предварительной приработки, когда в начальный период эксплуатации отказавшие, бракованные элементы заменяются новыми. Сформированная так партия элементов улучшает свои характеристики надежности.

2. Опасность отказов с увеличением времени эксплуатации уменьшается. Элемент в процессе эксплуатации ухудшает свои характеристики надежности. Элемент «стареет». Это очевидный, наиболее распространенный случай.

3. Опасность отказов в процессе эксплуатации не меняется, остается постоянной. Элемент в процессе эксплуатации «не стареет», остается «новым». Очевидно, что это некая теоретическая модель.

Рассмотрим подробнее последний случай. Если $\lambda(t) = Const$, то согласно (2.10) получим

$$P(t) = e^{-\int_0^t \lambda dt} = e^{-\lambda t} \quad (3.1)$$

Полученный закон надежности называется экспоненциальным и имеет большое значение в теории надежности. Частота отказов в соответствии с (2.3) равна

$$q(t) = -\frac{dP}{dt} = -\lambda e^{-\lambda t} \quad (3.2)$$

Средняя наработка на отказ согласно (2.11) равна

$$T_0 = \int_0^\infty P(t)dt = \int_0^\infty e^{-\lambda t}dt = -\frac{e^{-\lambda t}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda} \quad (3.3)$$

Вычислим вероятность надежной работы на интервале, равном среднему времени отказов $t = T_0$

$$\mathbf{P(T_0) = e^{-\lambda T_0} = e^{-\frac{\lambda}{\lambda}} = e^{-1} \approx 0.3}$$

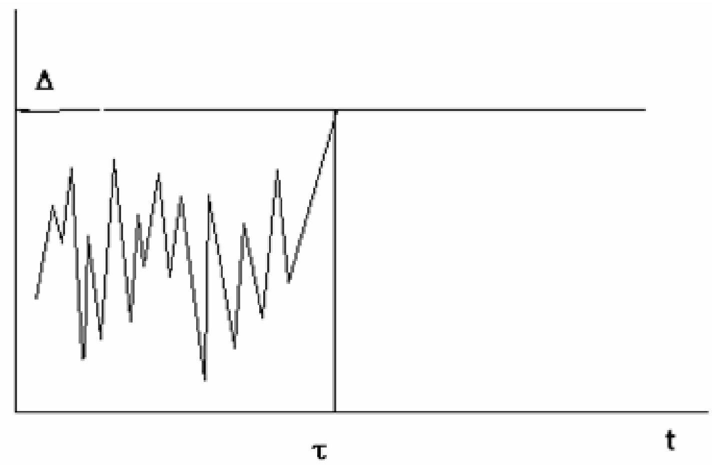
Надежность очень мала, примерно только треть элементов останется работоспособными. Поэтому для ответственных элементов время эксплуатации следует выбирать много меньше среднего времени жизни элемента ($\mathbf{t << T_0}$). Тогда вероятность безотказной работы согласно (3.1) примерно равна

$$\mathbf{P(t) = t^{-\frac{t}{T_0}} \approx 1 - \frac{t}{T_0}} \quad (3.4)$$

а вероятность отказа примерно равна

$$\mathbf{Q(t) = \frac{t}{T_0}} \quad (3.5)$$

Вероятность отказа на малом интервале времени $\mathbf{\Delta t}$ у исправного элемента равна $\lambda\mathbf{\Delta t}$. Эта вероятность не зависит от проработанного времени, если элемент исправен, то его дальнейшее поведение не зависит от прошлой предыстории. Верно и обратное утверждение. Для моделей отказов, когда предыстория не влияет на поведение отказов, справедлив экспоненциальный закон надежности. Такие модели отказов называются схемой мгновенных повреждений.



Модель мгновенных повреждений
Рис. 3.1

Пусть на элемент воздействует случайная ударная перегрузка (рис3.1). Когда она меньше критической Δ, отказа нет, нет никаких мелких повреждений, ухудшающих надежностные характеристики. Элемент остается по всем показателям как «новый». Как только перегрузка превышает критический уровень, элемент отказывает. Если перегрузка представляет собой случайный процесс типа белого шума, все отсчеты которого независимы, предыстория эксплуатации на момент отказа не влияет. То есть такая модель отказа соответствует экспоненциальному закону надежности. Недостаток экспоненциального закона в том, что он отражает редко встречающую модель отказов. Однако далее будет показано, что только этот закон позволяет построить практически применимые, инженерные методы расчетов.

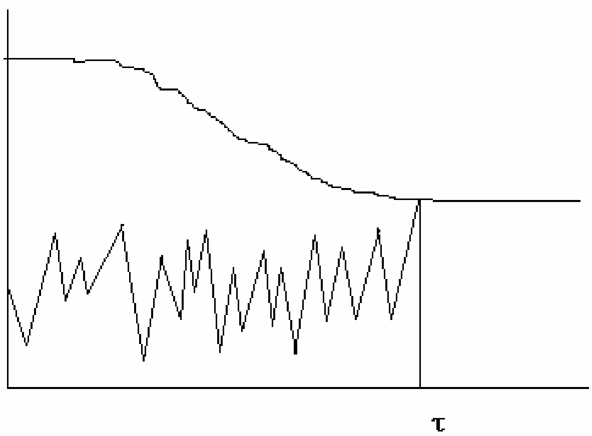


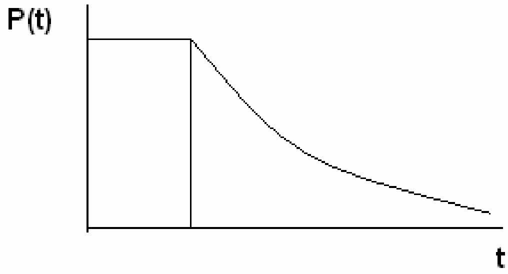
Схема с переменным порогом
Рис. 3.2

Рассмотрим еще одну схему мгновенных повреждений, но с переменным во времени порогом (Рис. 3.2). Эта схема отличается от схемы, приведенной на рис 3.1 тем, что порог – граница допустимых перегрузок переменный. Вначале он велик и перегрузка никогда не может его достигнуть. Затем порог снижается и в некоторый момент времени $\mathbf{t_0}$ фиксируется на некотором постоянном значении, которое достигается перегрузкой.

Имеем при $\mathbf{t < t_0}$ элемент не отказывает $\mathbf{P(t) = 1}$.

При $\mathbf{t > t_0}$ имеем экспоненциальный закон надежности $\mathbf{P(t) = e^{-\lambda(t-t_0)}}$

График вероятности безотказной работы показан на рис. 3.3

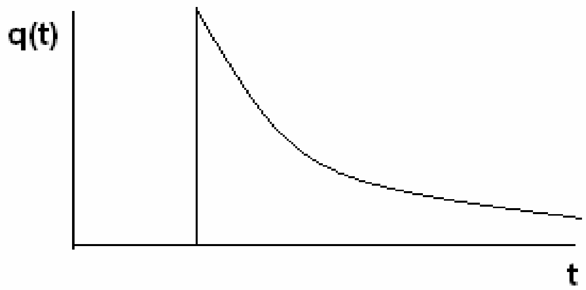


Закон с переменным порогом
Рис 3.3

Для плотности вероятности отказа имеем (см. рис. 3.4)

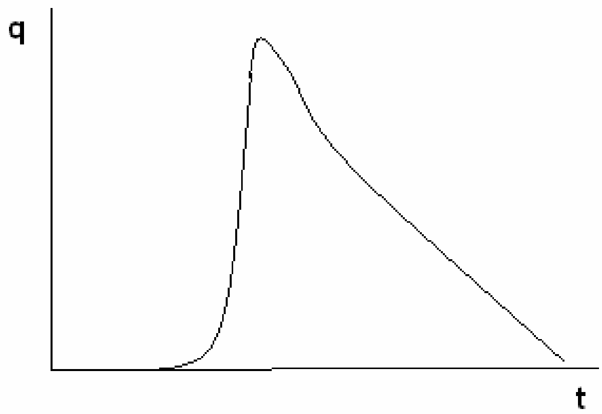
$$q(t) = -\frac{dP}{dt} \qquad \text{при } t < t_0 \qquad q(t) = 0$$

$$\text{при } t > t_0 \qquad q(t) = \lambda e^{-\lambda(t-t_0)}$$



Опасность отказа
Рис. 3.4

Эта модель может служить приближением для результатов экспериментального определения закона надежности, полученного в виде, показанного на рис. 3.5.



Экспериментальная кривая частоты отказов
Рис. 3.5

Контрольные вопросы

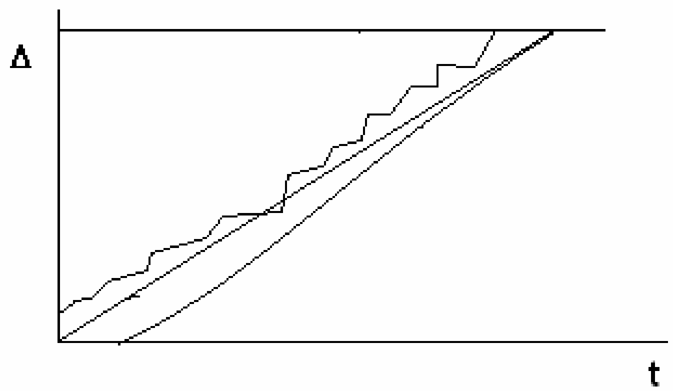
1. Что такое экспоненциальный закон надежности?

2 В каком соотношении находятся среднее время безотказной работы и реальное время эксплуатации?

3. Опишите модель мгновенных повреждений.

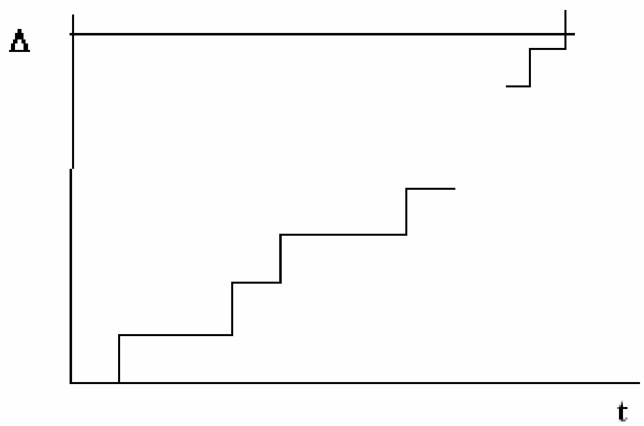
4. Модель накопления повреждений. Гамма – распределение времени безотказной работы

Модель накопления повреждений предполагает, что действие возмущений в процессе эксплуатации, износ элемента все время накапливается, увеличивается. Это накопление может начинаться от разных начальных условий из-за случайных мелких дефектов производства, может нарастать с разной скоростью из-за случайного характера воздействий и т. д. Предположим, что существует некоторый постоянный порог, по достижении которого происходит отказ (рис.4.1)



Модель накопления повреждений
Рис. 4.1

Рассмотрим следующую модель накопления повреждений (рис 4.2).



Принятая схема накопления повреждений
Рис. 4.2

В случайные независимые моменты времени скачком происходит увеличение повреждения. Все скачки равны по амплитуде. Примем величину скачка, равной $\mathbf{a}$. Примем также, что соотношение между скачком и порогом таково, что отказ произойдет после $\mathbf{r}$ скачков. Пусть времена скачков образуют простейший поток событий, то есть удовлетворяют следующим условиям: события независимы; поток стационарный; в один момент времени может произойти только один скачок. Известно, что для простейшего потока или потока Пуассона, вероятность наступления точно $\mathbf{k}$ событий за время $\mathbf{T}$ равна

$$\mathbf{P_k(T) = \frac{(\alpha T)^k}{k!} t^{-\alpha T}} \tag{4.1}$$

Здесь α средняя частота скачков.

Отказ наступит после некоторого целого числа скачков $\mathbf{r}$. Тогда $\mathbf{P(t)}$ - функция надежности есть вероятность того, что за время $\mathbf{t}$ произойдет меньше чем $\mathbf{r}$ скачков

$$\mathbf{P(t) = \sum_{k=o}^{k=r-1} P^k(t) = \sum_{k=0}^{k=r-1} \frac{(\alpha t)^k}{k!} e^{-\alpha t}} \tag{4.2}$$

Найдем плотность вероятности отказов для такого распределения

$$\begin{aligned}
\mathbf{q(t)} &= \mathbf{-\frac{dP}{dt} = -\frac{d}{dt}\{e^{-\alpha t}[1 + \frac{\alpha t}{1!} + \frac{(\alpha t)^2}{2!} + ... + \frac{(\alpha t)^{r-1}}{(r-1)!}]\} =} \\
&= \mathbf{\alpha e^{-\alpha t}[1 + \frac{\alpha t}{1!} + \frac{(\alpha t)^2}{2!} + ... + \frac{(\alpha t)^{r-1}}{(r-1)!}] - e^{\alpha t}[\alpha + \alpha\frac{\alpha t}{1!} + ... + \alpha\frac{(\alpha t)^{r-2}}{(r-2)!}] =} \\
&= \mathbf{\frac{1}{(r-1)!}\alpha^r t^{r-1} e^{-\alpha t}}
\end{aligned} \tag{4.3}$$

Получен двухпараметрический $(\alpha, \mathbf{r})$ закон надежности, которым можно было бы аппроксимировать экспериментальные характеристики. Но для этого надо распространить (4.3) на непрерывные значения $\mathbf{r}$. В принципе в (4.3) $\mathbf{r}$ может быть непрерывным всюду кроме аргумента факториала.

Существует гамма-функция, которая для целых положительных значений аргумента совпадает со значением факториала, и может быть принята как распространение понятия факториала на нецелые значения аргумента.

Выражение для гамма-функции есть

$$\mathbf{\Gamma(r)=\int_0^{\infty} x^{r-1}e^{-x}dx} \tag{4.4}$$

Для целых положительных $\mathbf{r}$ значение гамма-функции равно

$$\mathbf{\Gamma(r)=(r-1)!} \tag{4.5}$$

Заменяя в (4.3) факториал на гамма-функцию, получим

$$\mathbf{q(t)=\frac{1}{\Gamma(r)}\alpha^{r}t^{r-1}e^{-\alpha t}} \tag{4.6}$$

Тогда для непрерывных значений $\mathbf{r}$ вероятность отказов равна

$$\mathbf{Q(t)=\frac{\alpha^{r}}{\Gamma(r)}\int_0^{t}\tau^{r-1}e^{-\alpha\tau}d\tau}, \tag{4.7}$$

а функция надежности равна

$$\mathbf{P(t)=\frac{\alpha^{r}}{\Gamma(r)}\int_t^{\infty}\tau^{r-1}e^{-\alpha\tau}d\tau} \tag{4.8}$$

Для средней наработки на отказ имеем

$$\mathbf{T_0=\int_0^{\infty}tq(t)dt=\int_0^{\infty}\frac{\alpha^2\tau^2}{\Gamma(r)}e^{-\alpha\tau}d\tau=\frac{1}{\Gamma(r)}\int_0^{\infty}x^{r}e^{-x}\frac{dx}{\alpha}=\frac{\Gamma(r+1)}{\Gamma(r)\alpha}} \tag{4.9}$$

Здесь сделана подстановка $\mathbf{x}=\alpha\tau$, следовательно $\mathbf{d}\tau=\frac{\mathbf{dx}}{\alpha}$.

Для гамма-функции имеет место рекуррентное соотношение, которое следует из (4.5)

$$\mathbf{\Gamma(r+1)=r\Gamma(r)} \tag{4.10}$$

Тогда для среднего времени жизни из (4.9) имеем

$$\mathbf{T_0 = \frac{r}{\alpha}} \tag{4.11}$$

Обратим внимание, что полученный результат согласуется с результатами для экспоненциального закона. Для модели мгновенных повреждений достаточно одного события – достижения порога $\mathbf{(r=1)}$, которое не зависит от предыстории процесса, чтобы произошел отказ. Средняя частота событий в пуассоновском потоке α по вероятности равна величине, обратной среднему времени между отказами $\mathbf{\frac{1}{T_{0e}}}$ для экспоненциального закона надежности. Следовательно,

$$\alpha = \lambda_{\mathbf{e}}$$

Тогда для экспоненциального закона из (4.11) получаем известный нам результат

$$\mathbf{T_{0e} = \frac{1}{\lambda_e}}$$

Для дисперсии времени жизни имеем

$$\begin{aligned}\mathbf{D(\tau)} &= \int_0^\infty \tau^2 \mathbf{q(\tau)d}\tau - \mathbf{T_0}^2 = \int_0^\infty \frac{\alpha^{\mathbf{r}}\tau^{\mathbf{r+1}}}{\Gamma(\mathbf{r})}\mathbf{e}^{-\alpha\tau}\mathbf{d}\tau - (\frac{\mathbf{r}}{\alpha})^2 = \\ &= \frac{1}{\Gamma(\mathbf{r})}\int_0^\infty \mathbf{x^{r+1}e^{-x}}\frac{1}{\alpha^2}\mathbf{dx} - (\frac{\mathbf{r}}{\alpha})^2 = \frac{\Gamma(\mathbf{r+2})}{\alpha^2\Gamma(\mathbf{r})} - (\frac{\mathbf{r}}{\alpha})^2\end{aligned} \tag{4.12}$$

Здесь учтено, что согласно (4.10)

$$\mathbf{\Gamma(r+2) = (r+1)\Gamma(r+1) = (r+1)r\Gamma(r)}$$

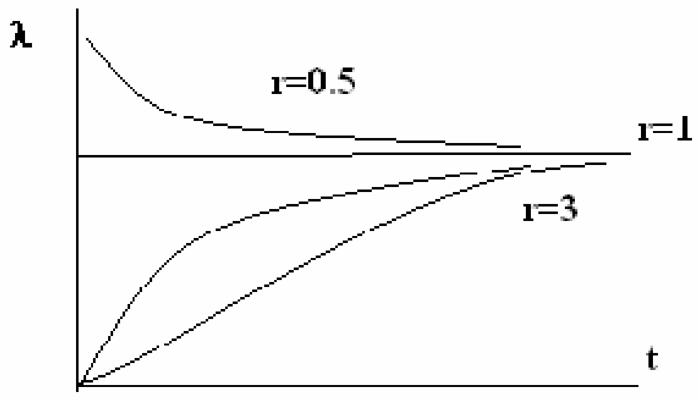
Тогда

$$\mathbf{D(\tau)} = \frac{\mathbf{(r+1)r - r^2}}{\alpha^2} = \frac{\mathbf{r}}{\alpha^2} \tag{4.13}$$

Опасность отказа равна

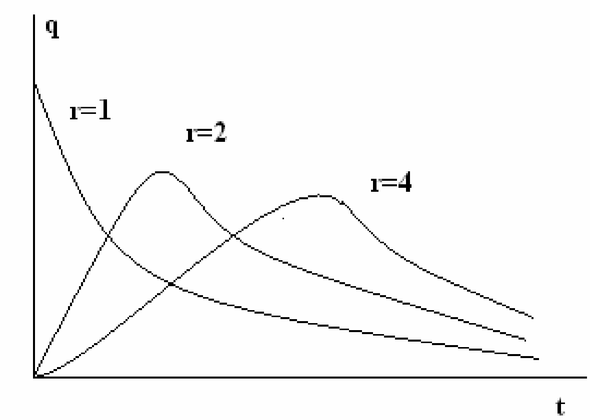
$$\lambda\mathbf{(t)} = -\frac{\mathbf{P'(t)}}{\mathbf{P(t)}} = \frac{\mathbf{q(t)}}{\mathbf{P(t)}}$$

Опасность отказа для этого закона надежности вычисляется сложным образом. Однако при малых временах можно принять $\mathbf{P(t) \approx 1; e^{-\alpha t} \approx 1}$, тогда из (4.6) следует, что зависимость опасности отказа от времени имеет вид $\lambda\mathbf{(t) \approx At^{r-1}}$, что позволяет оценить поведение опасности отказа на начальном интервале. Точные зависимости показаны на рис 4.3.



Зависимость опасности отказов от параметра
Рис. 4.3

При $\mathbf{r > 1}$ имеем модель стареющих элементов (опасность отказов увеличивается). При $\mathbf{r = 1}$ имеем экспоненциальный закон надежности. Это соответствует модели мгновенных повреждений. Одного скачка достаточно для отказа. При $\mathbf{r < 1}$ имеем убывающую опасность отказа. Частота отказов представлена на рис.4.4. В пределе при $\mathbf{r \to \infty}$ закон гамма- распределения стремится к нормальному закону надежности.



Плотность вероятности отказов -
- частота отказов
Рис. 4.4

Контрольные вопросы

1. Опишите модель повреждений закона гамма-распределения времени безотказной работы?

2. Ч то такое гамма- функция, зачем она используется при описании закона гамма-распределения?

3. Какой вид имеет зависимость опасности отказа от времени при законе гамма-распределения?

5. Нормальный закон надежности

Закон надежности вида гамма-распределения с увеличением параметра $\mathbf{r}$ приближается к нормальному закону

$$\mathbf{q(t)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi D}}} \mathbf{e}^{-\frac{\mathbf{(t-T_0)^2}}{\mathbf{2D}}}, \tag{5.1}$$

где $\mathbf{T_0} = \frac{\mathbf{r}}{\alpha}$ и $\mathbf{D} = \frac{\mathbf{r}}{\alpha^2}$.

Замена закона гамма-распределения нормальным законом может быть оправдана при $\frac{\mathbf{T_0}}{\sqrt{\mathbf{D}}} > \mathbf{3.5}$. Разница в функциях надежности при такой замене будет меньше 10% за исключением больших значений $\mathbf{t}$, где вероятности отказа малы.

Нормальное распределение в виде (5.1) дает значение $\mathbf{P(0) < 1}$, так как нормальное распределение имеет бесконечный диапазон изменения аргумента и, следуя (5.1), необходимо учитывать ненулевое значение плотности вероятности отказа при отрицательном аргументе.

Ситуацию необходимо корректировать, так как вся оценка надежности элемента происходит при временах относительно малых, по сравнению со средним временем жизни. С другой стороны точность каждого знака имеет огромное значение. Для того, чтобы элемент имел значение функции надежности в начальный период эксплуатации 0.99999 вместо 0.9999, проходило порядка десяти лет напряженной конструкторской и технологической работы.

В выражение (5.1) вводится некоторый поправочный коэффициент, такой чтобы значение функции надежности в нуле было бы равно единице.

Введем нормированную величину

$$\mathbf{x} = \frac{\mathbf{t - T_0}}{\sigma} \tag{5.2}$$

позволяющую использовать таблицы интеграла вероятности. Тогда

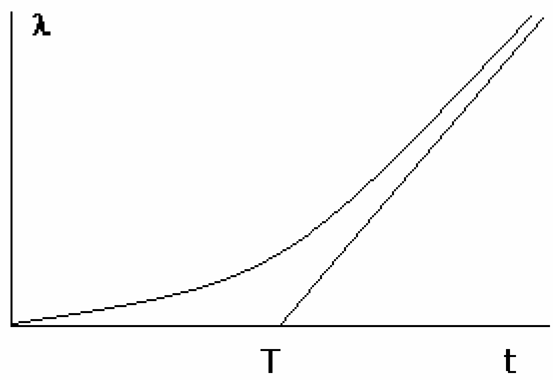
$$\mathbf{P(t)=\frac{1}{\sqrt{2\pi}}\int\limits_{x}^{\infty}e^{-\frac{x^2}{2}}dx} \tag{5.3}$$

$$\mathbf{P(0)=\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{T_0}{\sigma}}^{\infty}e^{-\frac{x^2}{2}}dx} \tag{5.4}$$

Функция надежности будет равна нулю, если в качестве нормального закона надежности примем следующую нормированную функцию

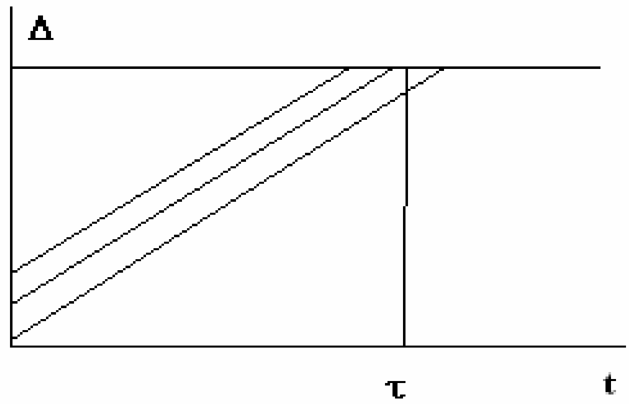
$$\mathbf{P_n(t)=\frac{P(t)}{P(0)}=\frac{\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{t-T_0}{\sigma}}^{\infty}e^{-\frac{x^2}{\sigma}}dx}{\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{-T_0}{\sigma}}^{\infty}e^{-\frac{x^2}{\sigma}}dx}} \tag{5.5}$$

Частота отказов как плотность вероятности будет представлять собой нормальную плотность вероятности, нормированную с коэффициентом $\mathbf{1/P(0)}$ и с математическим ожиданием, равным $\mathbf{T_0}$. Опасность отказов имеет вид, показанный на рис 5.1. Асимптота опасности отказов для больших времен имеет вид $\mathbf{y=\frac{t-T_0}{\sigma}}$.



Опасность отказа нормального закона
Рис. 5.1

Нормальному закону надежности можно сопоставить следующую модель отказов. Это накапливающиеся повреждения со случайными начальными условиями, распределенными по нормальному закону. Скорость накопления детерминированная, одинаковая для всех случаев (рис 5.2). Тогда момент отказа τ также будет нормальной случайной величиной. Эта модель хороша для описания надежности элементов массового хорошо налаженного производства.



Накопление повреждений при нормальном законе

Рис. 5.2

6. Закон Вейбулла – Гнеденко

Закон надежности Вейбулла – Гнеденко имеет место при следующей модели накопления повреждений. На элемент действуют не одно а несколько воздействий, вызывающих отказ. Каждое воздействие может действовать по своей схеме накопления повреждений (мгновенное повреждение или накопление повреждений и т. д.). Воздействуют они на разные параметры элемента (механическая прочность, прочность пайки, истирание, износ, сопротивление изоляции и т. д.). Все воздействия взаимно независимы и для каждого существует свое время жизни $\tau_{\mathbf{i}}$. Отказ элемента наступит при наименьшем значении времени жизни из набора $\tau_{\mathbf{i}}$, то есть для того воздействия, которое раньше других превысит порог. Случайным является значение $\tau_{\mathbf{i}}$, случайным является воздействий, вызывающее отказ. Следовательно, для описания такой модели отказа необходимо найти закон распределения случайной величины τ, которое равно наименьшему значению из набора случайных величин $\tau_{\mathbf{i}}$

$$\tau = \min(\tau_1, \tau_2, ...\tau_n) \qquad (6.1)$$

Пусть вероятность отказа для каждого воздействия или интегральный закон распределения каждого $\tau_{\mathbf{i}}$ при малом времени имеют одинаковый вид

$$\mathbf{Q(t) \approx ct^{\alpha}},$$

где $\alpha > \mathbf{0}$. Тогда для случайной величины τ получена следующая функция надежности

$$\mathbf{P(t) = e^{-\gamma t^{\alpha}}} \qquad (6.2)$$

Для плотности вероятности отказа имеем

$$\mathbf{q(t)} = -\frac{\mathbf{dP}}{\mathbf{dt}} = \gamma\alpha\mathbf{t}^{\alpha-1}\mathbf{e}^{-\gamma t^{\alpha}} \qquad (6.3)$$

Для опасности отказа имеем

$$\lambda = \frac{\mathbf{q(t)}}{\mathbf{P(t)}} = \frac{\gamma\alpha \mathbf{t}^{\alpha-1}\mathbf{e}^{-\gamma t^{\alpha}}}{\mathbf{e}^{-\gamma t^{\alpha}}} \tag{6.4}$$

Средняя наработка на отказ получена в виде

$$\mathbf{T_0} = \frac{\Gamma(\frac{1}{\alpha}+1)}{\gamma^{\frac{1}{\alpha}}} \tag{6.5}$$

Этот закон получил распространение, поскольку он сохраняет некоторые положительные свойства экспоненциального закона надежности. С помощью двух параметров α и γ можно имитировать различные модели отказов. Параметр α позволяет получить модели «стареющих» и «молодеющих элементов». При $\alpha = 1$ имеем экспоненциальный закон надежности. Параметр γ позволяет выбирать масштабы.

7. Учет условий эксплуатации

По сведениям, указанным в [3], учет условий эксплуатации при расчете надежности может быть произведен по следующей методике. На участке нормальной эксплуатации расчет опасности отказов каждого элемента проводится по формуле

$$\lambda_{\mathbf{r}} = \lambda_{\mathbf{0}}(\prod_{\mathbf{i}} \mathbf{K}_{\mathbf{i}})$$

Здесь $\mathbf{K}_{\mathbf{i}}$ - поправочный коэффициент по $\mathbf{i}$ -му фактору, учитывающий значение электрической нагрузки, климата, других условий эксплуатации, которые вы учитываете в соответствии с ТЗ. Среди коэффициентов $\mathbf{K}_{\mathbf{i}}$ есть общие для всех элементов и специальные.

Общие коэффициенты:

$\mathbf{К}_{\mathbf{н}}$ - коэффициент нагрузки (учитывает электрическую нагрузку и температуру среды),

$\mathbf{К}_{\mathbf{э}}$ - коэффициент эксплуатации (учитывает особенности эксплуатации),

$\mathbf{К}_{\mathbf{пр}}$ - коэффициент приемки (учитывает степень жесткости требований к контролю качества и правил приемки),

$\mathbf{К}_{\mathbf{рн}}$ - коэффициент роста надежности (учитывает предполагаемый рост надежности за счет различных мероприятий в процессе эксплуатации).

Коэффициент эксплуатации упрощено можно представить в виде

$$\mathbf{К}_{\mathbf{э}} = \mathbf{К}_{\mathbf{м}}\mathbf{К}_{\mathbf{вл}}\mathbf{К}_{\mathbf{д}}$$

где

$\mathbf{К}_{\mathbf{м}}$ - влияние механических нагрузок, вибраций, ударов,

$\mathbf{К}_{\mathbf{вл}}$ - влияние влажности,

$\mathbf{К}_{\mathbf{д}}$ - влияние пониженного давления.

Специальные коэффициенты для интегральных микросхем, например, учитывают:

- сложность микросхемы,

- колебания питающего напряжения,

- тип корпуса,

- степень освоения технологии изготовления.

Для полупроводниковых приборов специальные коэффициенты учитывают:

- функциональное назначение,

- максимально допустимую нагрузку по мощности,

- диапазон рабочего напряжения,

- частота и мощность импульса,

- тип корпуса.

Для коммутационных элементов учитывается:

- число задействованных контактов,

- количество коммутаций в час.

8. Надежность систем с последовательным соединением

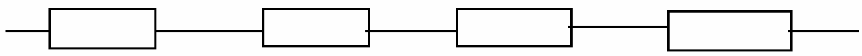
После того как рассмотрены вопросы надежности элементов, перейдем к рассмотрению вопросов надежности систем. Предполагается, что система состоит из элементов, причем известна структура системы и характер её работы. Под структурой системы будем понимать следующее:

1. Известна надежность всех элементов по отдельности.
2. Для каждого элемента известно, вызывает ли его отказ отказ всей системы.
3. Влияет ли отказ элемента на характеристики надежности других элементов, в том числе на их отказ.

В этом разделе предполагается, что элементы независимы в смысле отказа.

Рассмотрим случай, когда элементы в системе соединены последовательно в смысле надежности. Это означает, что отказ любого элемента приводит к отказу всей системы.

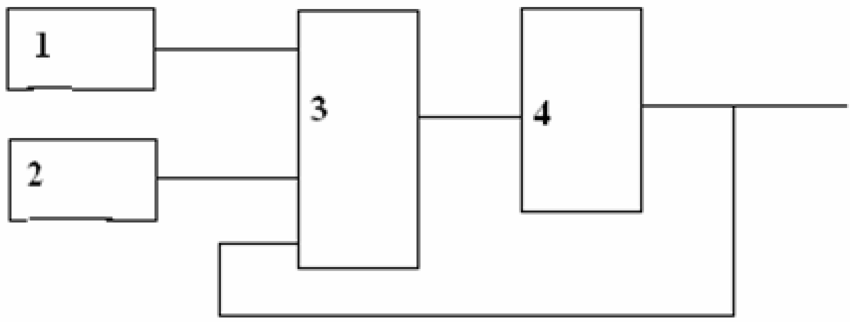
Пусть имеется цепочка сопротивлений, соединенных в электрическую схему последовательно (рис. 8.1). Здесь отказ любого сопротивления выводит из строя всю цепь, следовательно, сопротивления включены в схему последовательно в смысле надежности.



Последовательное соединение сопротивлений
Рис. 8.1

Рассмотрим упрощенную схему системы стабилизации (рис. 8.2). Здесь

1. Гироскоп.
2. Датчик скорости.
3. ПИД – регулятор.
4. Управляемый объект.



Система стабилизации

Рис. 8.2

Отказ любого элемента в этой схеме приведет к отказу системы. Поэтому в этой схеме элементы соединены последовательно в смысле надежности, хотя их функциональное соединение образует систему с отрицательной обратной связью. То есть соединения в смысле надежности и функциональные схемы могут не совпадать.

Для безотказной работы системы с последовательным соединением в течении времени $\mathbf{t}$ необходимо, чтобы все элементы работали безотказно в течение этого времени. Так как по предположению элементы независимы в смысле отказов, следовательно, вероятность безотказной работы системы есть произведение вероятностей безотказной работы всех элементов

$$\mathbf{P_c(t) = P_1(t)P_2(t)...P_n(t)} \quad , \tag{8.1}$$

где

$\mathbf{P_c(t)}$ - функция надежности системы,

$\mathbf{P_i(t)}$ - функция надежности элемента.

Функция надежности имеет вид

$$\mathbf{P(t) = e^{-\int_0^t \lambda(t)dt}} \tag{8.2}$$

Тогда

$$\mathbf{e^{-\int_0^t \lambda_c(t)dt} = e^{-\int_0^t \lambda_1(t)dt}\, e^{-\int_0^t \lambda_2(t)dt} \ldots\ldots\, e^{-\int_0^t \lambda_n(t)dt}} \tag{8.3}$$

Следовательно, при любом законе надежности элементов опасность отказа системы равна

$$\lambda_c(t) = \sum_i \lambda_i(t) \qquad (8.4)$$

Если у всех элементов экспоненциальный закон надежности, то все их опасности отказов постоянны, следовательно, опасность отказа системы постоянна

$$\lambda_c(t) = Const$$

следовательно, система имеет тоже экспоненциальный закон надежности. Это условие сохранения вида закона надежности соблюдается только для экспоненциального закона надежности. В этом причина, по которой инженерные схемы расчета надежности основываются на применении экспоненциального закона.

Средняя наработка на отказ для экспоненциального закона равна

$$T_c = \frac{1}{\lambda_c}$$

$$T_i = \frac{1}{\lambda_i}$$

Тогда

$$\frac{1}{T_c} = \frac{1}{T_1} + \frac{1}{T_2} + \dots + \frac{1}{T_n}$$

Или

$$T_c = \frac{1}{\frac{1}{T_1} + \frac{1}{T_2} + \dots + \frac{1}{T_n}} \qquad (8.3)$$

Если все элементы в смысле надежности одинаковы, получим

$$T_c = \frac{T_i}{n} \qquad (8.4)$$

Среднее время жизни системы элемента уменьшилось обратно пропорционально числу элементов. Чем больше элементов в системе с параллельным соединением, тем хуже надежность системы.

Пусть система должна иметь среднее время жизни

$$\mathbf{T_0 = 1год}$$

Пусть в системе всего 100 одинаковых элементов. Тогда каждый элемент должен иметь среднее время жизни 100 лет. В этом опасность увеличения числа элементов при последовательном соединении.

Рассмотрим влияние этого эффекта на возможности экспериментального определения характеристик надежности. Рассмотрим следующую задачу.

Система состоит из n последовательно соединенных элементов. Необходимо экспериментально подтвердить, что при $\mathbf{t = t_0}$ функция надежности равна

$$\mathbf{P(t_0) = P^*.}$$

Найдем число элементов, которые надо поставить на испытание. Для экспоненциального закона

$$\mathbf{P^* = P_c(t_0) = e^{-\lambda n t_0} \approx 1 - \lambda n t_0}$$

Вероятность отказа системы равна

$$\mathbf{Q^* = \lambda n t_0}$$

Пусть поставлено на испытание $\mathbf{k}$ элементов. Тогда среднее число отказавших ко времени $\mathbf{t_0}$ элементов равно

$$\mathbf{N \approx \lambda k t_0}$$

Чтобы получить статистический результат с приемлемым доверительным интервалом, надо получить порядка $\mathbf{N = 300...500}$ подтверждений. Составим отношение

$$\frac{Q^*}{N} = \frac{n\lambda t_0}{k\lambda t_0} = \frac{n}{k}$$

Отсюда число испытываемых элементов равно

$$k = \frac{nN}{1 - P^*}$$

Даже при очень низкой требуемой надежности $P^* = 0.99$ (для ответственных систем число девяток после запятой порядка пяти – шести) получаем нереальную величину $k = 10^6$.

Поэтому для экспериментального подтверждения характеристик надежности применяют так называемые ускоренные и утяжеленные испытания. В этих испытаниях условия эксплуатации на несколько порядков хуже обычных, опасности отказов поэтому на несколько порядков выше и необходимое число отказов набирается за приемлемое время. Затем результаты ускоренных испытаний экстраполируются на нормальные условия эксплуатации. Сомнения здесь уместны.

Поставим задачу определения вероятности отказа системы $Q_c(t)$ по вероятностям отказа элементов $Q_i(t)$. Имеем

$$Q_c(t) = 1 - P_c(t) =$$
$$= 1 - \prod_{i=1}^{n} P_i(t) = 1 - \prod_{i=1}^{n} [1 - Q_i(t)] = \sum_{i=1}^{n} Q_i(t) + \sum_{i}^{n} \sum_{j}^{n} Q_i(t) Q_j(t)$$

Если

$$\sum_i Q_i(t) << 1$$

(требуется высокая надежность системы), можно положить

$$Q_c(t) \cong \sum_i Q_i(t) , \qquad (8.5)$$

причем погрешность не превосходит величины

$$\frac{1}{2}\sum_{i=1}^{n} Q_i^2(t)$$

Вероятности отказов складываются. С ростом числа элементов надежность системы с последовательным соединением падает.

Контрольные вопросы

1. Что такое последовательное соединение в смысле надежности?

2 Чему равна вероятность отказа системы при последовательном соединении элементов?

3. Чему равна опасность отказа системы при последовательном соединении элементов?

4 Какой закон надежности имеет система при последовательном соединении элементов, если закон надежности элементов – экспоненциальный?

9. Системы с параллельным соединением

Элементы в системе соединены параллельно в смысле надежности, если отказ системы наступает после отказа всех элементов, входящих в систему. Например, при параллельном электросоединении сопротивлений проводимость такой цепи будет равна нулю только тогда, когда откажут все сопротивления. Здесь понятия функционального соединения и соединения в смысле надежности совпадают. Рассмотрим некоторую комплексную навигационную систему самолета, в которой измерение координат ведется несколькими подсистемами: гироскопическими, инерциальными, доплеровскими, радиолокационными. Если при отказе отдельных устройств продолжается вычисление координат, которое прекращается только при отказе всех подсистем, нужно считать, что подсистемы соединены параллельно в смысле надежности.

Так как отказ системы с параллельным соединением наступает после отказа всех элементов, вероятность отказа системы $\mathbf{Q_c(t)}$ есть вероятность произведения событий – отказ отдельного элемента. Для элементов, независимых в смысле надежности, вероятность произведения событий равна произведению вероятности. Поэтому

$$\mathbf{Q_c(t) = \prod_{i=1}^{n} Q_i(t)} \tag{9.1}$$

где $\mathbf{Q_i(t)}$ есть вероятность отказа i – го элемента.

Из (9.1) следует, что вероятность отказа системы уменьшается с ростом числа элементов, так как $\mathbf{Q_i(t) < 1}$.

Если все элементы системы одинаковы в смысле надежности

$$\mathbf{Q_c(t) = Q_i(t)^n}. \tag{9.2}$$

Функция надежности системы равна при этом

$$\mathbf{P_c(t) = 1 - Q_i(t)^n} \tag{9.3}$$

Вычислим среднее время жизни системы при экспоненциальном законе надежности элементов

$$T_0 = \int_0^{\infty} P_c(t)dt = \int_0^{\infty} [1 - (1 - e^{-\lambda t})^n]dt$$

Сделаем замену переменных

$$x = 1 - e^{-\lambda t}$$

Тогда

$$e^{-\lambda t} = 1 - x;$$

$$t = -\frac{1}{\lambda}\ln(1 - x);$$

$$dt = \frac{1}{\lambda}\frac{1}{1 - x}dx$$

Проводя дальнейшие вычисления, получим

$$T_0 =$$

$$= \frac{1}{\lambda}\int_0^1 \frac{1 - x^n}{1 - x}dx = \frac{1}{\lambda}\int_0^1 (1 + x + x^2 + ... + x^{n-1})dx = \frac{1}{\lambda}[x + \frac{x^2}{2} + \frac{x^3}{3} + ... + \frac{x^n}{n}]|_0^1$$

или

$$T_0 = \frac{1}{\lambda}(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n})$$

Если n велико,

$$T_0 = \frac{1}{\lambda}[\ln(x) + C]$$

где C=0.577 – постоянная Бернулли.

Видно, что при параллельном соединении элементов среднее время жизни системы растет, то есть увеличивается надежность системы.

Найдем вероятность надежной работы системы – функцию надежности при экспоненциальном законе надежности элементов

$$P_c(t) = 1 - \prod_{i=1}^{n} Q_i(t) = 1 - \prod_{i=1}^{n} (1 - e^{-\lambda_i t})$$

В системах с высокой надежностью $\lambda_i t << 1$, тогда

$$P_c(t) \cong 1 - \prod_{i=1}^{n} [1 - 1 + \lambda_i t + O(\lambda_i t)^2$$

где $O(\lambda_i t)^2$ полином второго порядка малости.

В результате получаем функцию надежности системы при параллельном соединении элементов в виде

$$P_c(t) = 1 - \prod_{i=1}^{n} (\lambda_i t) \qquad (9.4)$$

Например, для системы из двух параллельно соединенных элементов имеем

$$P_c(t) = 1 - \lambda_1 \lambda_2 t^2 \qquad (9.5)$$

Отметим принципиальную особенность систем с параллельным соединением. Надежность таких систем не только выше, но производная по времени функции надежности в начальный момент эксплуатации нулевая. Надежность на этом отрезке практически абсолютная, функция надежности равна единице, вероятность отказа равна нулю.

Пусть имеются элементы с опасностью отказа, равной $\lambda = 10^{-4} 1/\text{час}$. За 100 часов работы вероятность отказа элемента равна $\lambda t = 0.01$. Вероятность безотказной работы равна

$$P(100) = 1 - \lambda t = 0.99$$

Рассмотрим систему из двух одинаковых параллельно соединенных элементов. Вероятность безотказной работы для нее, согласно (9.5) равна

$$P_c(100) = 1 - (\lambda t)^2 = 0.9999$$

Включение второго параллельно соединенного элемента эквивалентно понижению λ характеристики элемента на два порядка, что достигается годами конструкторской и технологической работы, которая не всегда заканчивается успешно.

Поэтому часто с целью повышения надежности в систему искусственно включают на некоторых участках параллельно элементы, подсистемы, узлы. Такое изменение структуры системы называется дублированием или резервированием.

Схемы резервирования могут быть различными.

Горячий резерв. При этом виде резерва элементы включены параллельно и все включаются одновременно: и основной, и резервные. Поэтому принимают, что при горячем резерве характеристики надежности и основного элемента и резервных элементов все время одинаковы.

Холодный резерв. При этом виде резерва сначала включается основной элемент, после его отказа он заменяется резервным элементом, затем после отказа этого элемента включается следующий резервный и т. д. Здесь принимается, что элемент, находящийся в резерве отказать не может.

Облегченный резерв. Облегченный резерв является разновидностью горячего резерва, при котором принимается, что у элемента, находящегося в резерве, λ - характеристика меньше, чем у основного элемента. После того как резервный элемент будет подключен как основной, его $\lambda -$ характеристика увеличится.

Проектирование резервированных систем предполагает разработку: устройств определения исправности основного и резервных элементов (нельзя включить в работу неисправный резервный элемент, как правило, это катастрофа); устройств подключения резервных элементов в основную схему.

Контрольные вопросы.

1. Дайте определение соединения, параллельного в смысле надежности.

2. Как меняется надежность системы с параллельным соединением элементов?

3. Нарисуйте примерное поведение λ - характеристики при облегченном резерве

10. Виды резервирования [4]

Показано выше, что введение в схему дополнительно параллельно включенного элемента повышает надежность системы. Такое структурное изменение схемы называется резервированием. В общем случае резервированием называется способ повышения надежности системы за счет использования дополнительных средстви возможностей, которые являются избыточными для выполнения функций системы (без учета надежности). То есть в систему помимо средств, достаточных для выполнения задач, проводятся дополнительные мероприятия с целью повышения надежности. Поэтому в настоящее время вместо термина «резервирование» используют часто термин «внесение избыточности».

Виды и методы резервирования разнообразны.

Структурное резервирование. Это способ повышения надежности, состоящий в применении дополнительных элементов, устройств, которые не являются необходимыми для выполнения функций системы. Они используются (включаются) только тогда, когда возникает опасность отказа или произошел отказ. В идеальной, абсолютно надежной системе они не нужны.

Резервирование называется **общим**, если заменяется вся последовательная цепь, **раздельным**, если заменяются отдельные элементы, **групповым**, если заменяется часть элементов.

Скользящее резервирование не закрепляет резервный элемент (группу) за определенным участком, а замещает им любой отказавший участок.

Кратностью резервирования называется отношение

$$\mathbf{q} = \frac{\mathbf{m}}{\mathbf{k}}$$

где

$\mathbf{m}$ - число резервных элементов (групп),

$\mathbf{k}$ - число основных элементов (групп).

Резервные элементы повышают вес, габариты, энергопотребление в системе. Они требуют диагностических устройств, которые бы обеспечивали подключение только исправных резервных элементов взамен отказавших. Сами устройства подключения не так просты.

Функциональное резервирование. При таком способе резервирования надежность системы повышается за счет устранения последствий отказа перераспределением функций между элементами системы, как это делает живой организм. Лишних, дополнительных элементов при таком способе резервирования в системе нет. Некоторые элементы временно начинают дополнительно выполнять необходимые системе функции, иногда ухудшая качество выполнения собственных функций.

Временное резервирование (резервирование времени). В этом случае у системы заранее создаются некоторые запасы продукта, ресурса, который система в случае отказа может некоторое время расходовать, выполняя свои функции и восстанавливать свои технические характеристики.

Резерв времени может создаваться за счет:

- уменьшения оперативного времени, отводимого на выполнение задания,

- накопления запасов продукции на складах,

- накопления заявок в БЗУ,

- накопления энергии в аккумуляторах, воды в водохранилищах и т. д.

Для систем с резервом времени отказ системы возникает в случае, когда после отказа неисправность не удается устранить за время, равное резерву времени.

Резерв времени может быть пополняемым и непополняемым. Мгновенно пополняемый резерв восстанавливается скачком сразу после окончания ремонта и включения элемента в работу. В общем случае закон пополнения резерва должен быть задан.

Информационное резервирование. Примером создания информационного резервирования является информационная избыточность такая, как дублирование передаваемых данных, увеличение разрядности помехоустойчивого кода, позволяющего бороться с помехами, сбоями и т. п. Архивные, журнальные файлы, откаты в системах управления базами данных.

Алгоритмическое резервирование. Алгоритмическое резервирование это введение избыточности в алгоритмы обработки информации минимальной сложности: дублирование ввода данных, двойной просчет, контрольные тестовые точки и т. п.

Контрольные вопросы.

1. Перечислите виды резервирования.
2. Дополните раздел примерами информационного резервирования.
3. Что такое кратность резервирования?

11. Марковские процессы с дискретными состояниями и непрерывным временем [2]

В этом и ряде последующих разделов основным методом расчета надежности систем является теория дискретных Марковских систем с непрерывным временем. Теория допускает простые, инженерные методы расчета. Показаны возможности применения этой теории при расчете надежности резервированных систем и систем с зависимыми отказами. Рассматривается надежность систем с восстанавливаемыми элементами. Ввиду идентичности используемого теоретического аппарата рассматриваются примеры систем массового обслуживания, в том числе расчет буферных запоминающих устройств.

Напомним, что марковским случайный процесс будет, если для каждого момента времени $\mathbf{t_0}$ вероятность любого состояния системы при $\mathbf{t \geq t_0}$ зависит только от текущего состояния при $\mathbf{t = t_0}$ и не зависит от всех предыдущих состояний системы при $\mathbf{t < t_0}$.

Случайный процесс называется процессом с дискретными состояниями, если система $\mathbf{S}$ может принимать конечное число возможных состояний: $\mathbf{S_1, S_2, ... S_n}$, а сам процесс состоит в том, что время от времени система $\mathbf{S}$ мгновенно переходит из одного состояния в другое.

Случайный процесс называется процессом с дискретным временем, если смена состояний возможна в строго определенные, заранее фиксированные, известные моменты времени – «шаги»: $\mathbf{t_1, t_2, ...}$ Для каждого шага должны существовать вероятности перехода системы из любого состояния в любое $\mathbf{p_{ij}}$. В общем случае эти вероятности перехода зависят от номера шага, состояния, откуда идет переход, состояния, куда идет переход и состояний, принятых системой на предыдущих шагах. Для марковских процессов не зависят от предыдущих состояний. Марковский процесс называется однородным, если вероятности перехода не зависят также от номера шага.

Тогда для системы с $\mathbf{n}$ состояниями должны быть заданы $\mathbf{n^2}$ вероятностей перехода. Вероятности нахождения системы в $\mathbf{k}$-ом состоянии на $\mathbf{m}$-ом шаге $\mathbf{P_k{}^m}$ вычисляются по формуле полной вероятности

$$\mathbf{P_k{}^m = \sum_{i=1}^{n} p_{ik} P_i{}^{m-1}} \tag{11.1}$$

Здесь

$\mathbf{n}$ - число возможных состояний,

$\mathbf{P_i{}^{m-1}}$ - вероятность $\mathbf{i}$ – го состояния на предыдущем шаге.

Рекуррентный процесс (11.1) начинается с $\mathbf{m=0}$ по заданным в соответствии с начальными условиями вероятностям $\mathbf{P_i{}^0}$.

На практике чаще встречаются ситуации, когда смена состояний происходит не в фиксированные, а в случайные моменты времени. Если эти моменты времени – непрерывные случайные величины, то процесс называется непрерывной цепью Маркова или дискретным марковским процессом с непрерывным временем.

Для непрерывных цепей Маркова не существуют определенные ранее вероятности перехода для фиксированного шага, поскольку вероятность любого фиксированного момента времени равна нулю. Поэтому и вероятность перехода в точно заданный момент времени равна нулю.

Вводится условная плотность вероятности перехода

$$\lambda_{\mathbf{ij}} = \mathbf{lim}_{\mathbf{\Delta t \to 0}} \frac{\mathbf{p_{ij}(\Delta t)}}{\mathbf{\Delta t}} \tag{11.2}$$

где

$\mathbf{p_{ij}(\Delta t)}$ - вероятность перехода за ближайший интервал времени $\mathbf{\Delta t}$, которая для малого интервала согласно (11.2) равна

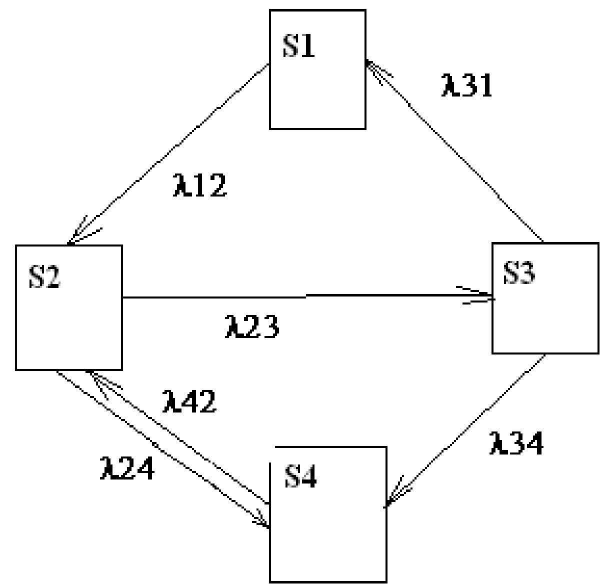
$$\mathbf{p_{ij}(\Delta t)} \approx \lambda_{\mathbf{ij}} \mathbf{\Delta t}, \tag{11.3}$$

если (11.2) существует.

Когда λ_{ij} не зависят от времени, процесс называется однородным.

Покажем метод определения вероятностей состояний системы на следующем примере.

Система **S** имеет четыре возможных состояния S_1, S_2, S_3, S_4. Размеченный граф состояний системы показан на рис. 11.1, где стрелками обозначены возможные переходы. Из-за технических недостатков используемого графического редактора индексы на рисунке представлены строчными числами.



Граф состояний системы
Рис. 11.1

Найдем значение $P_1(t)$- вероятность того, что система в момент времени **t** находится в состоянии S_1. Сначала найдем вероятность того, что система находится в состоянии S_1 в момент времени $t + \Delta t$, то есть $P_1(t + \Delta t)$.

Попасть в состояние S_1 система согласно графу на рис.11.1 может только двумя способами:

1. В момент **t** система была в состоянии S_1, а за время Δt из этого состояния не вышла. Условная вероятность выхода из этого состояния (которое может быть только в состояние S_2) равна $\lambda_{12}\Delta t$. А условная

вероятность невыхода из состояния S_1 равна $1-\lambda_{12}\Delta t$. Следовательно, полная вероятность рассматриваемого случая

$$P_1(t)[1-\lambda_{12}\Delta t]. \tag{11.4}$$

2. В момент t система была в состоянии S_3, а за интервал времени Δt перешла в состояние S_1. Условная вероятность этого случая равна $\lambda_{31}\Delta t$, а полная вероятность этого случая есть

$$P_{31}(t)\,\lambda_{31}(t)\Delta t \tag{11.5}$$

Вероятность других переходов в состояние S_1 за интервал времени Δt имеет второй порядок малости $(\Delta t)^2$ или еще меньше, так как эти переходы проходят через несколько состояний. Например, переход из S_2 через S_3 в S_1 имеет вероятность, равную $\lambda_{23}\lambda_{31}(\Delta t)^2$.

Следовательно, вероятность нахождения системы в состоянии S_1 в момент времени $t+\Delta t$ есть вероятность суммы двух независимых событий и равна

$$P_1(t+\Delta t)=P_1(t)[1-\lambda_{12}\Delta t]+P_{31}(t)\,\lambda_{31}(t)\Delta t$$

Переносим $P_1(t)$ в левую часть и делим обе части равенства на Δt. Получим

$$\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t}=-\lambda_{12}P_1(t)+\lambda_{31}P_3(t)$$

При $\Delta t\to 0$, переходя к пределу, получаем первое дифференциальное уравнение, которому удовлетворяет функция $P_1(t)$

$$\dot{P}_1(t)=-\lambda_{12}P_1(t)+\lambda_{31}P_3(t) \tag{11.6}$$

Рассмотрим состояние S_2 и найдем вероятность этого состояния в момент времени $t+\Delta t$, то есть вероятность $P_2(t+\Delta t)$. В состояние S_2 система может попасть следующими способами

1. Из состояния S_1, вероятность этого равна

$$P_1 \lambda_{12} \Delta t \tag{11.7}$$

2. Из состояния S_4, вероятность этого случая равна

$$P_4 \lambda_{42} \Delta t \tag{11.8}$$

3. Оставшись в состоянии S_2, вероятность этого случая равна

$$P_2(t)[1 - (\lambda_{23}\Delta t + \lambda_{24}\Delta t)] \tag{11.9}$$

Следовательно,

$$P_2(t + \Delta t) = P_1\lambda_{12}\Delta t + P_4\lambda_{42}\Delta t + P_2(t)[1 - (\lambda_{23}\Delta t + \lambda_{24}\Delta t)]$$

Аналогично (11.6), получим

$$\dot{P}_2(t) = \lambda_{12}P_1(t) + \lambda_{42}P_4(t) - (\lambda_{23} + \lambda_{24})P_2(t) \tag{11.10}$$

Обратим внимание на структуру уравнений (11.6) и (11.10). В левой части находится производная вероятности состояния, а в правой части имеется столько слагаемых, сколько стрелок графа связано с этим состоянием. Если стрелка направлена из состояния, слагаемое берется со знаком минус. Если стрелка направлена в состояние, слагаемое имеет знак плюс. Каждое слагаемое равно произведению условной плотности вероятности перехода, соответствующее данной стрелке, умноженной на вероятность состояния, из которого исходит стрелка. Таким образом сформулированы правила составления дифференциальных уравнений для определения вероятностей состояний непрерывных цепей Маркова, называемых уравнениями Колмогорова для непрерывных цепей Маркова.

Пользуясь этими правилами запишем остальные уравнения для состояний S_3 и S_4.

$$\dot{P}_3(t) = \lambda_{23}P_2(t) - (\lambda_{31} + \lambda_{34})P_3(t) \tag{11.11}$$

$$\dot{P}_4(t) = \lambda_{24}P_2(t) + \lambda_{34}P_3(t) - \lambda_{42}P_4(t) \tag{11.12}$$

Интегрирование системы дифференциальных уравнений (11.6), (11.10)-(11.12) определяет вероятности состояний как функции времени. Начальные условия определяются по начальному состоянию системы. Например, если при **t=0** система находилась в состоянии $\mathbf{S_1}$, то начальные условия равны

$$
\begin{aligned}
&\mathbf{P_1(0) = 1;}\\
&\mathbf{P_2(0) = 0;}\\
&\mathbf{P_3(0) = 0;}\\
&\mathbf{P_4(0) = 0.}
\end{aligned}
$$

Если в начальный момент система находится в неопределенном состоянии, необходимо задавать начальные вероятности состояний.

Некоторое упрощение при интегрировании системы уравнений можно получить, используя вместо одного из дифференциальных уравнений соотношение

$$\sum_{\mathbf{i}} \mathbf{P_i(t) = 1} \tag{11.13}$$

Это очевидное соотношение удовлетворяет системе. Действительно, по структуре системы уравнений сумма всех правых частей системы уравнений равна нулю. Следовательно,

$$\sum_{\mathbf{i}} \mathbf{\dot{P}_i(t)} = \frac{\mathbf{d}}{\mathbf{dt}} \sum_{\mathbf{i}} \mathbf{P_i(t) = 0}$$

Учитывая начальные условия, получаем (11.13), использование которого позволяет понизить порядок линейной системы дифференциальных уравнений.

Необходимо помнить, что уравнения Колмогорова можно использовать только для марковских систем, так как в процессе вывода уравнений условные плотности вероятности перехода $\lambda_{\mathbf{ij}} = \lambda_{\mathbf{ij}}\mathbf{(t)}$ зависели только от текущего момента времени и не зависели от предыдущего поведения процесса.

Контрольные вопросы

1. Что такое в цепях Маркова: дискретные состояния? Что такое непрерывное время?

2. Каковы правила составления уравнений Колмогорова?

3. Как можно понизить порядок системы дифференциальных уравнений Колмогорова?

12. Применение теории непрерывных цепей Маркова к расчету надежности резервированных систем

Уравнения Колмогорова могут быть всегда использованы для расчета надежности, если у элементов имеет место экспоненциальный закон надежности. При других законах надежности применимость уравнений Колмогорова требует исследования.

Граф состояний при отказе элемента

Рис. 12.1

Рассмотрим граф состояний одного невосстанавливаемого элемента при возможном отказе, представленном на рис 12.1. Состояние $\mathbf{S_0}$ элемент исправен. При отказе элемент переходит в состояние $\mathbf{S_1}$. Условной плотностью вероятности перехода из $\mathbf{S_0}$ в $\mathbf{S_1}$ является лямбда – характеристика, которая является условной плотностью вероятности отказа (переход в состояние $\mathbf{S_1}$) при условии, что элемент был исправен (вначале находился в состоянии $\mathbf{S_0}$). Если у элемента экспоненциальный закон надежности, опасность отказа постоянна и не зависит от времени. $\lambda(\mathbf{t}) = \lambda = \mathbf{Const}$. Процесс безусловно марковский.

Рассмотрим резервированную систему, которая состоит из одного основного элемента и трех резервированных элементов, находящихся в «облегченном « резерве. Все элементы имеют экспоненциальный закон надежности с опасностями отказов, равными:

λ_1 - для основного элемента,

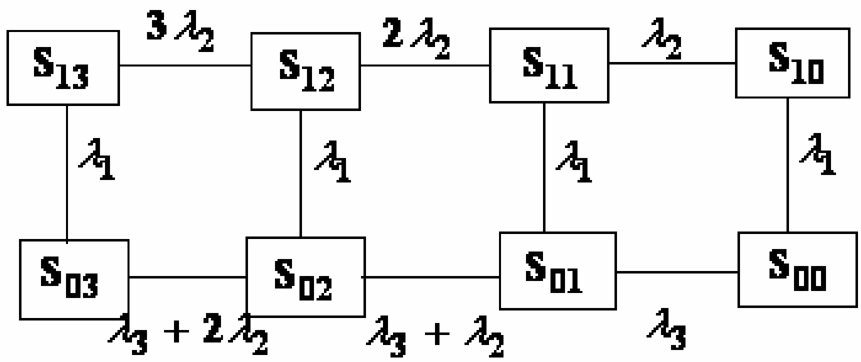
λ_2 - для резервного элемента в режиме резерва,

λ_3 - для резервного элемента в режиме работы.

Обозначим состояния системы через **Sij**, где **i=0,1** – число исправных основных элементов; **j=0,1,2,3** – число исправных резервных элементов.

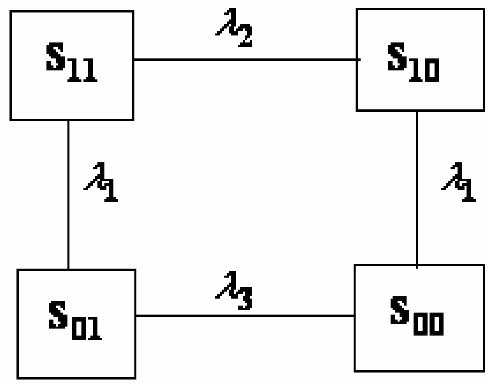
Смена состояний происходит в момент отказа какого-либо элемента. Поэтому условные плотности вероятности смены состояний являются опасностями отказов.

В случае экспоненциальных законов надежности опасности отказов λ не зависят от времени. Условные плотности вероятности переходов зависят только от состояния, в котором находится система, и не зависят от времени пребывания в этом состоянии. В этом случае для расчета надежности могут быть использованы уравнения Колмогорова. Граф состояний системы показан на рис. 12.2.



Граф отказов резервированной системы
Рис. 12.2

Переход системы из состояния $\mathbf{S_{1n}}$ в состояние $\mathbf{S_{1(n-1)}}$ происходит при отказе одного из N резервных элементов, находящихся в режиме резерва. Условная плотность вероятности перехода при этом равна сумме вероятностей, то есть $\mathbf{n} * \lambda_2$. Переход из состояния $\mathbf{S_{1n}}$ в состояние $\mathbf{S_{0n}}$ происходит при отказе основного элемента, поэтому условная плотность вероятности перехода равна λ_1. Переход из состояния $\mathbf{S_{0n}}$ в состояние $\mathbf{S_{0(n-1)}}$ происходит при отказе одного из N резервных элементов. Но основной элемент уже неисправен и заменен одним из резервных элементов. То есть один из резервных элементов находится в режиме работы и условная плотность вероятности перехода равна $\lambda_3 + (\mathbf{n} - 1) * \lambda_2$. Состояние $\mathbf{S_{00}}$ – отказ системы.



Система с одним резервным элементом

Рис. 12.3

Расчет проведем для системы с одним резервным элементом. Граф состояний показан на рис 12.3. По графу состояний записываем систему уравнений Колмогорова

$$\dot{\mathbf{P}}_{11}(\mathbf{t}) = -(\lambda_1 + \lambda_2)\mathbf{P}_{11} \qquad (12.1)$$
$$\dot{\mathbf{P}}_{01}(\mathbf{t}) = \lambda_1\mathbf{P}_{11} - \lambda_3\mathbf{P}_{01} \qquad (12.2)$$
$$\dot{\mathbf{P}}_{10}(\mathbf{t}) = \lambda_2\mathbf{P}_{11} - \lambda_1\mathbf{P}_{10} \qquad (12.3)$$
$$\dot{\mathbf{P}}_{00}(\mathbf{t}) = \lambda_3\mathbf{P}_{01} + \lambda_1\mathbf{P}_{10} \qquad (12.4)$$

которую необходимо решать при начальных условиях

$$\mathbf{P}_{11}(0) = 1$$
$$\mathbf{P}_{01}(0) = \mathbf{P}_{10}(0) = \mathbf{P}_{00}(0) = 0$$

(все элементы системы в начальный момент исправны и система находится в состоянии $\mathbf{S_{11}}$).

Для определения функции надежности системы, равной

$$\mathbf{P_c}(\mathbf{t}) = \mathbf{P}_{11}(\mathbf{t}) + \mathbf{P}_{01}(\mathbf{t}) + \mathbf{P}_{10}(\mathbf{t}) \qquad (12.5)$$

из (12.1) для заданных условий получаем

$$\mathbf{P}_{11}(\mathbf{t}) = \mathbf{e}^{-(\lambda_1 + \lambda_2)\mathbf{t}} \qquad (12.6)$$

Подставляя (12.6) в (12.2) и (12.3), получим

$$\dot{P}_{01} + \lambda_3 P_{01} = \lambda_1 e^{-(\lambda_1+\lambda_2)t}$$
$$\dot{P}_{10} + \lambda_1 P_{10} = \lambda_2 e^{-(\lambda_1+\lambda_2)t}$$

Решения этих уравнений при нулевых начальных условиях равны соответственно

$$P_{01}(t) = \frac{\lambda_1}{\lambda_1+\lambda_2-\lambda_3}[e^{-\lambda_3 t} - e^{-(\lambda_1+\lambda_2)t}]$$
$$P_{10}(t) = e^{-\lambda_1 t} - e^{-(\lambda_1+\lambda_2)t} \qquad (12.7)$$

Подставляя (12.6) и (12.7) в (12.5), для функции надежности системы имеем

$$P_c(t) = \frac{\lambda_1}{\lambda_1+\lambda_2-\lambda_3}[e^{-\lambda_3 t} - e^{-(\lambda_1+\lambda_2)t}] + e^{-\lambda_1 t} \qquad (12.8)$$

Для среднего времени жизни системы имеем

$$T_{0c} = \int_0^{\infty} P_c(t)dt = \frac{\lambda_1}{\lambda_1+\lambda_2-\lambda_3}\left(\frac{1}{\lambda_3} - \frac{1}{\lambda_1+\lambda_2}\right) + \frac{1}{\lambda_1} =$$

$$= \frac{\lambda_1}{\lambda_3(\lambda_1+\lambda_2)} + \frac{1}{\lambda_1} = T_1 + T_3\frac{T_2}{T_1+T_2} \qquad (12.9)$$

где

$T_1 = \dfrac{1}{\lambda_1}$ - среднее время жизни основного элемента при экспоненциальном законе надежности,

$T_2 = \dfrac{1}{\lambda_2}$ - среднее время жизни резервного элемента в режиме резерва,

$T_3 = \dfrac{1}{\lambda_3}$ - среднее время жизни резервного элемента в режиме работы.

Рассмотрим свойства функции надежности системы на интервале времени, когда вероятность отказа элементов мала

$$\lambda_1 t << 1; \quad \lambda_2 t << 1; \quad \lambda_3 t << 1.$$

Тогда с точностью до малых третьего порядка выражение (12.9) примет вид

$$P_c(t) \cong \frac{\lambda_1}{\lambda_1+\lambda_2-\lambda_3}[1-\lambda_3 t+\frac{(\lambda_3 t)^2}{2}-...-1+(\lambda_1+\lambda_2)t-\frac{(\lambda_1+\lambda_2)^2 t^2}{2}+$$
$$+...]+1-\lambda_1 t+\frac{(\lambda_1 t)^2}{2}-...=1-\lambda_1(\lambda_2+\lambda_3)\frac{t^2}{2}+... \quad (12.10)$$

Вероятность отказа резервированной системы равна

$$Q_c(t) \cong \lambda_1(\lambda_2+\lambda_3)\frac{t^2}{2} \quad (12.11)$$

Положим

$$\lambda_1 = \lambda_3 = \lambda$$
$$\lambda_2 = k\lambda$$

Тогда

К=0 соответствует «холодному» резерву,

К=1 соответствует «горячему» резерву,

0<k<1 соответствует «облегченному» резерву.

В таблице 12.1 приведены вероятности отказа и средние времена жизни для различных случаев резерва. Видны преимущества «холодного» резерва. Видно, что для любого вида резерва имеет место нулевая первая производная вероятности отказа в начальный период эксплуатации.

Таблица 12.1

	$Q_c(t)$	T_{oc}
Нерезервированный элемент	λt	T_0
«Горячий» резерв	$(\lambda t)^2$	$1.5T_0$
«Облегченный» резерв	$\frac{1+k}{2}(\lambda t)^2$	$\frac{2+k}{1+k}T_0$
«Холодный» резерв	$\frac{1}{2}(\lambda t)^2$	$2T_0$

Контрольные вопросы

1. Что позволяет использовать уравнения Колмогорова для решения задач о надежности?

2. Каковы преимущества «облегченного» резерва перед «горячим» резервом?

13. Характеристики надежности и эксплуатации мгновенно восстанавливаемого элемента

До настоящего времени в конспекте рассматривался такой процесс работы систем и элементов, при котором после первого отказа элемента или системы эксплуатация их прекращалась. Однако, в основном большинство изделий после отказа, выхода из строя подвергаются ремонту, их функции восстанавливаются, и эксплуатация их продолжается.

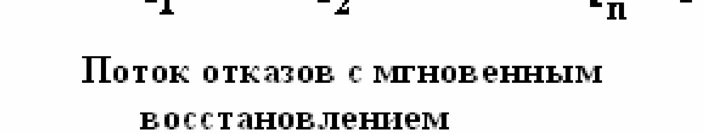
В этом разделе рассматривается случай, когда ремонт или замена отказавшего элемента на новый длится такое короткое время, что можно пренебречь временем ремонта и считать, что элемент после отказа восстанавливает все свои функции мгновенно. Тогда можно считать, что элемент все время находится в работоспособном состоянии, и отказов у него нет. Поэтому нас интересуют в этом процессе другие явления, связанные с обеспечением ремонта отказавших элементов:

- как отказы распределяются по времени, это необходимо для планирования работы ремонтных бригад;

- как распределяются по времени количество эксплуатируемых и ремонтируемых изделий, это необходимо для определения общего количества потребных изделий;

- как распределяется по времени количество потребных для ремонта деталей и др.

Рассмотрим вначале следующую простейшую модель. Элемент начинает работать в момент времени $t = 0$ (см. рис.13.1). Через случайный интервал времени



Поток отказов с мгновенным восстановлением

Рис.13.1

τ_1 в момент времени $t_1 = \tau_1$ элемент отказывает. Будем считать, что мы мгновенно нашли отказ и, либо мгновенно заменили элемент новым, либо мгновенно отремонтировали его так, что он полностью восстановил свои функции – стал «новым» и включили его в работу. Через случайный интервал времени τ_2 в момент времени $t_2 = \tau_1 + \tau_2$ элемент опять отказал и снова мгновенно восстановился.

Предположим, что случайные величины τ_i независимы и имеют одинаковые интегральные законы распределения

$$Q(t) = P(\tau_i < t),$$

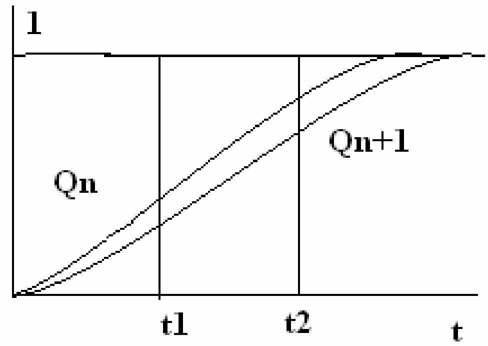
которые являются вероятностями отказа, вычисленными в соответствии с теорией надежности. Моменты времен отказов t_i образуют случайный поток событий, который мы назовем процессом восстановления.

Найдем вероятность того, что за время t произойдет n отказов или больше. Число отказов за заданное время есть случайная величина. Обозначим её как $\nu(t)$. Событие $\nu(t) \geq n$ равносильно такому событию, когда время t_n (время n - го отказа) находится внутри рассматриваемого промежутка времени t. Следовательно, закон распределения $\nu(t)$ есть

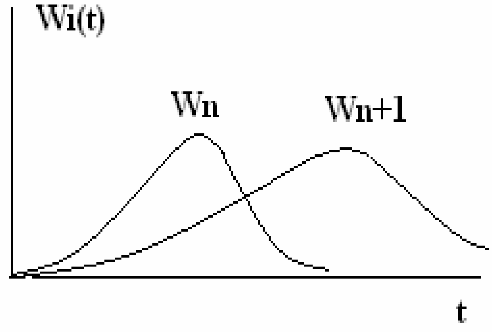
$$Q_n t) = [\nu(t) \geq n] = P(t_n < t) = P(\tau_1 + \tau_2 + ... + \tau_n \leq t) \qquad (13.1)$$

Задача кажется простой. Закон распределения суммы большого числа независимых случайных величин известен. Это нормальное распределение, все параметры которого известны, если n задано. Нам же необходимо найти закон распределения числа отказов.

Предположим, что интегральный закон распределения суммы n случайных величин - $Q_n(t)$ и $n+1$ - ой случайной величины $Q_{n+1}(t)$ известны (рис.13.2)

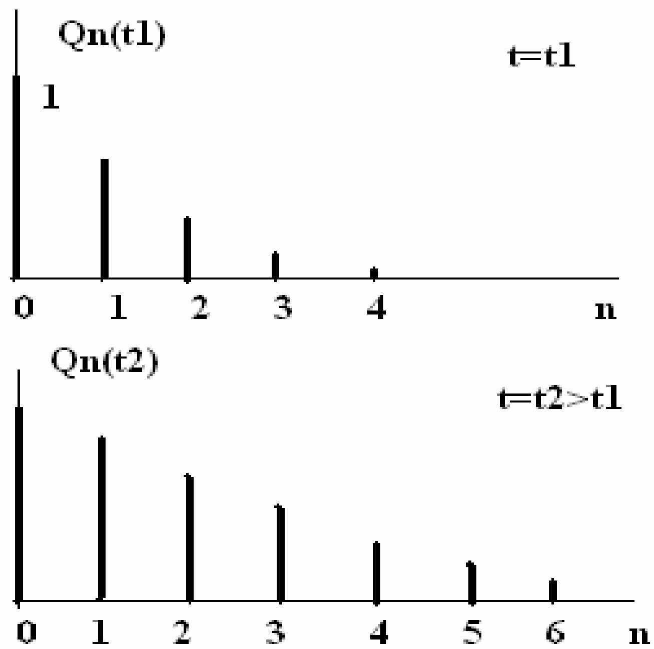


Интегральные законы распределения
Рис. 13.2



Плотности вероятности распределения числа отказов
Рис.13.3

На рис. 13.3 представлены плотности вероятности этих сумм. Эти законы распределения вероятностей могут быть найдены по известным в теории вероятности способам. Как отмечалось ранее, если **n** достаточно велико, то закон распределения суммы независимых величин приближается к нормальному и строится легко.



Интегральный закон для числа отказов
Рис. 13.4

Однако для определения параметров процесса восстановления необходимо найти закон распределения числа отказов **n**. Он может быть найден, например, как сечение кривых на рис 13.2 для различных моментов времени. Примерный вид распределения показан на рис. 13.4. Напоминаем, что значения $\mathbf{Q_n(t) = P[\nu(t) > n]}$. Так как вероятность того, что число отказов больше нуля для любого времени, равно единице, имеем $\mathbf{Q_0(t) = 1}$.

Вероятность того, что в системе произойдет ровно $\mathbf{n}$ отказов - $\mathbf{P_n(t)}$, учитывая, что случайная величина $\nu(\mathbf{t})$ целое число, можно записать в виде

$$\mathbf{P_n(t) = P[\nu(t) = n] = P[n+1 > \nu(t) \geq n] = Q_n(t) - Q_{n+1}(t)} \qquad (13.2)$$

Например, можно на рис. 13.4 для каждого момента времени построить график разностей соседних отсчетов.

Большую роль при изучении процесса восстановления имеет математическое ожидание числа отказов одного, произошедших от начала работы до момента времени $\mathbf{t}$, которое называется функцией восстановления $\mathbf{H(t)}$. По определению

$$\mathbf{H(t) = M[\nu(t)] = \sum_{n=1}^{\infty} n * P_n(t)} \qquad (13.3)$$

Используя (13.2), получим

$$\mathbf{H(t) = \sum_{n=1}^{\infty} n[Q_n(t) - Q_{n+1}(t)] = Q_1 + 2Q_2 + 3Q_3 ... - Q_2 - 2Q_3 - 3Q_4 ... =}$$

$$\mathbf{= \sum_{n=1}^{\infty} Q_n(t)} \qquad (13.4)$$

Вместо функции восстановления часто рассматривают дифференциальную характеристику

$$\mathbf{h(t) = \frac{dH(t)}{dt} = \sum_{n=1}^{\infty} \frac{dQ_n(t)}{dt} = \sum_{n=1}^{\infty} w_n(t)} \qquad (13.5)$$

которая называется плотностью восстановления или интенсивностью потока отказов. Она определяет среднее число отказов $\mathbf{\Delta\overline{n}(t)}$ на малом интервале восстановления $\mathbf{t ... t + \Delta t}$

$$\Delta\overline{\mathbf{n}}(\mathbf{t}) = \mathbf{h}(\mathbf{t})\Delta\mathbf{t}\mathbf{N_0} \qquad (13.6)$$

где $\mathbf{N_0}$ - количество элементов в начальный момент.

Несмотря на кажущуюся одинаковость, плотность восстановления отличается от частоты отказов. Частота отказов $\mathbf{q(t)}$ определяет среднее число отказов $\Delta\mathbf{n(t)}$ на этом же интервале $\mathbf{t...t} + \Delta\mathbf{t}$

$$\Delta\mathbf{n}(\mathbf{t}) = \mathbf{q}(\mathbf{t})\Delta\mathbf{t}\mathbf{N_0} \qquad (13.7)$$

для партии из $\mathbf{N_0}$ невосстанавливаемых элементов. Плотность восстановления есть характеристика процесса, когда все вышедшие из строя элементы восстанавливаются.

Найдем зависимость между частотой отказов и плотностью восстановления. В общем количестве отказавших элементов, определяемом по (13.6), часть элементов $\Delta\mathbf{n(t)}$ (см. (13.7)) окажется из партии, работающей с начального момента без отказов. Остальная часть $\Delta\mathbf{n_2(t)}$ включает в себя элементы, которые ранее отказывали и были восстановлены. Запишем

$$\Delta\overline{\mathbf{n}}(\mathbf{t}) = \Delta\mathbf{n}(\mathbf{t}) + \Delta\mathbf{n_2}(\mathbf{t}) \qquad (13.8)$$

На интервале $\tau\mathbf{...}\tau + \Delta\tau$, где $\tau < \mathbf{t}$ выйдет из строя (впервые или после промежуточного восстановления)

$$\Delta\mathbf{n_2}(\tau) = \mathbf{h}(\tau)\Delta\tau\mathbf{N_0} \qquad (13.9)$$

элементов. Из этого числа после восстановления на интервале $\mathbf{t...t} + \Delta\mathbf{t}$ первый раз откажут

$$\Delta\mathbf{n}_{2\tau}(\mathbf{t}) = \mathbf{h}(\tau)\Delta\tau\mathbf{N_0}\mathbf{q}(\mathbf{t} - \tau)\Delta\mathbf{t} \qquad (13.10)$$

элементов. Те элементы, которые еще раз откажут на промежуточном интервале $\tau_1\mathbf{...}\tau_1 + \Delta\tau_1$, где $\tau < \tau_1 < \mathbf{t}$ войдут в общее число элементов, отказавших на этом интервале и равное

$$\Delta\mathbf{n_2}(\tau_1) = \mathbf{h}(\tau_1)\Delta\tau_1\mathbf{N_0}.$$

Тогда для определения $\Delta\mathbf{n_2(t)}$ просуммировать (13.10) по всем интервалам $\Delta\tau$

$$\Delta\mathbf{n_2(t) = N_0 \Delta t \int_0^t h(\tau) q(t-\tau) d}\tau \tag{13.11}$$

Подставляя (13.6), (13.7) и (13.11) в (13.8) и, сокращая на $\mathbf{N_0 \Delta t}$, получим

$$\mathbf{h(t) = q(t) + \int_0^t h(\tau) q(t-\tau) d}\tau \tag{13.12}$$

Полученное соотношение позволяет по экспериментальным значениям $\mathbf{h(t)}$, полученным в процессе длительной эксплуатации, определить частоту отказа $\mathbf{q(t)}$, а затем и опасность отказа $\lambda\mathbf{(t)}$.

В некоторых случаях интегральное уравнение (13.12) может быть решено в операторной форме. Учитывая, что преобразование Лапласа для свертки равно

$$\mathbf{L[\int_0^t h(\tau) q(t-\tau) d\tau] = h(s) q(s)},$$

из (13.12) получаем

$$\mathbf{h(s) = \frac{q(s)}{1 - q(s)}}$$

или

$$\mathbf{q(s) = \frac{h(s)}{1 + h(s)}} \tag{13.13}$$

Отметим одно из свойств плотности восстановления

$$\mathbf{\lim_{t\to\infty} h(t) = \frac{1}{T_0}} \tag{13.14}$$

С течением времени процесс становится стационарным, а плотность восстановления становится обратной среднему времени жизни элемента.

Значит, для экспоненциального закона плотность восстановления становится равной опасности отказа

$$\lim h(t)_{t\to\infty} = \lambda(t) \qquad (13.15)$$

Это можно пояснить следующим образом. При экспоненциальном законе времена отказов в процессе восстановления образуют простейший пуассоновскийпоток событий,так как модель мгновенных повреждений, соответствующая экспоненциальному закону, предполагает, что время отказа не зависит от предыстории работы элемента. При этом законе средняя частота событий равна $\frac{1}{T} = \lambda$. С другой стороны средняя частота событий в процессе восстановления есть плотность восстановления.

Для экспоненциального закона, образующего простейший поток в процессе восстановления известна вероятность появления ровно n отказов за время t. Это формула Пуассона

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

В случае длительных периодов эксплуатации целесообразно использовать ассимптотическое распределение числа отказов, к которому стремится распределение отказов при любом законе надежности.

Вернемся к равенству (13.1)

$$Q_n(t) = P[\nu \ge n] = P[\tau_1 + \tau_2 + ... + \tau_n < t] \qquad (13.16)$$

Так как величины τ_i независимы и одинаково распределены с математическим ожиданием T_0 и дисперсией σ^2, то дробь

$$\delta_n = \frac{\tau_1 + \tau_2 + ... + \tau_n - nT_0}{\sigma\sqrt{n}} \qquad (13.17)$$

сходится при $n \to \infty$ к нормальной случайной величине с нулевым математическим ожиданием и единичной дисперсией. При том число отказов n будет близко по вероятности к величине $\frac{t}{T_0}$, поэтому положим

$$\mathbf{n} = \frac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{z_n}\sqrt{\mathbf{t}} \tag{13.18}$$

Смысл первого слагаемого пояснен выше. Второе слагаемое отражает тот факт, что $\mathbf{n}$ может быть любым целым положительным числом. Поэтому $\mathbf{z_n}$ принимает такие значения, чтобы это выполнялось. То, что второе слагаемое по вероятности мало, мы отметили множителем $\sqrt{\mathbf{t}}$.

Тогда первое неравенство в (13.16), вероятность которого определяется, можно преобразовать следующим образом. Вместо

$$\nu \geq \mathbf{n} = \frac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{z_n}\sqrt{\mathbf{t}}$$

запишем

$$\frac{\nu - \dfrac{\mathbf{t}}{\mathbf{T_0}}}{\sqrt{\mathbf{t}}} \geq \mathbf{z_n} \tag{13.19}$$

Вместо второго неравенства в (13.16)

$$\tau_1 + \tau_2 + \ldots + \tau_n < \mathbf{t},$$

учитывая (13.17), получим

$$\delta_n < \frac{\mathbf{t} - \mathbf{T_0}}{\sigma\sqrt{\mathbf{n}}}$$

Подставляя сюда (13.18), имеем

$$\delta_n < \frac{\mathbf{t} - (\dfrac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{z_n}\sqrt{\mathbf{t}})\mathbf{T_0}}{\sigma\sqrt{\dfrac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{z_n}\sqrt{\mathbf{t}}}} = -\frac{\mathbf{z_n}\sqrt{\mathbf{t}}\mathbf{T_0}}{\sigma\sqrt{\dfrac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{z_n}\sqrt{\mathbf{t}}}} \tag{13.20}$$

Используем в (13.16) преобразованные неравенства (13.19) и (13.20). Учтем при этом, что в (13.20) слагаемое $\mathbf{z_n}\sqrt{\mathbf{t}}$ по вероятности мало по сравнению с

$\frac{t}{T_0}$. Поэтому, когда вычисляют вероятности от неравенств, этим слагаемым можно пренебречь. Получим

$$\lim_{t\to\infty} P[\frac{\nu - \frac{t}{T_0}}{\sqrt{t}} \geq z_n] = P[\delta_n < -\frac{z_n T_0^{3/2}}{\sigma}]$$

Обозначим $x = \frac{z_n T_0^{3/2}}{\sigma}$. Тогда

$$\lim_{t\to\infty} P[\frac{\nu - \frac{t}{T_0}}{\frac{\sigma\sqrt{t}}{T^{3/2}}} \geq x] = P[\delta_n < -x]$$

Величина δ_n имеет нормальное распределение с нулевым математическим ожиданием и единичной дисперсией. Нормальное распределение симметрично. Поэтому такое же распределение как δ_n имеет величина $\frac{\nu(t) - \frac{t}{T_0}}{\frac{\sigma\sqrt{t}}{T_0^{3/2}}}$

Отсюда следует, что случайное число отказов $\nu(t)$ имеет нормальное распределение, математическое ожидание, равное

$$M_\nu = \frac{t}{T_0} \tag{13.21}$$

и дисперсию, равную

$$D_\nu = \frac{\sigma^2 t}{T_0^3} \tag{13.22}$$

Полученные соотношения позволяют найти простые оценки для определения числа запасных элементов.

Пусть, например, среднее время жизни элемента равно $\mathbf{T_0 = 100 часов}$. Дисперсия времени жизни равна $\sigma^2 = \mathbf{3600\ \ часов^2}$. Требуется с достоверностью 0.95 оценить число запасных элементов, необходимых для $\mathbf{t = 8000\ часов}$ эксплуатации.

Среднее число запасных элементов по (13.21) равно

$$\mathbf{M}_\nu = \frac{\mathbf{8000}}{\mathbf{100}} = \mathbf{80} \quad \textbf{элементов}$$

Но при расчете на это количество недостаточность запаса обнаружится в 50% случаев. Учитывая, что в данном случае оценка должна охватывать случаи потребления запасных элементов от нуля до исчерпания запаса, оценка вычисляется из условия (числовые значения выбираются из таблицы интеграла вероятности)

$$\frac{\mathbf{1}}{\sqrt{\mathbf{2}\pi}} \int_{\mathbf{0}}^{\mathbf{x_3}} \mathbf{t}^{-\frac{\mathbf{u^2}}{\mathbf{2}}} \mathbf{du} = \mathbf{0.95}$$

Определяем $\mathbf{x_3 = 1.65}$ и потребное количество запасных элементов равно

$$\nu = \mathbf{M}_\nu + \mathbf{x_3}\sqrt{\mathbf{D}_\nu} = \frac{\mathbf{t}}{\mathbf{T_0}} + \mathbf{x_3}\frac{\sigma\sqrt{\mathbf{t}}}{\mathbf{T_0}^{\mathbf{3/2}}} = \mathbf{80 + 9 = 89} \quad \textbf{элементов}$$

Недостаточность такого запаса обнаружится только в 5% случаев. Заметим, что с увеличением достоверности, то есть при стремлении лучше застраховаться, увеличивается число запасных элементов и теперь возникает угроза того, что запасные элементы остаются неиспользованными. Для вероятности обеспеченности, равной единице, число запасных элементов должно быть равным бесконечности.

Контрольные вопросы

1. Дайте определение процессу с мгновенным восстановлением.

2. Как по интегральному закону распределения фиксированного числа отказов по времени найти распределение числа отказов по времени?

3. Что такое плотность восстановления?

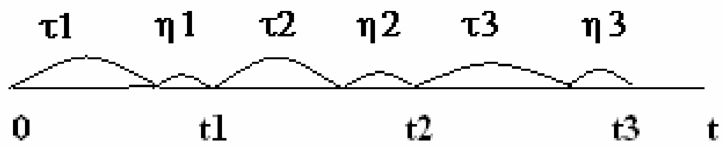
4. Каково соотношение между плотностью восстановления и частотой отказов?

5. Напишите выражения для математического ожидания и дисперсии асимптотического распределения числа отказов? Каков закон асимптотического распределения числа отказов?

6. В чем опасности заниженного или завышенного значения достоверности определения числа запасных элементов?

14. Процесс с конечным случайным временем восстановления

В отличие от предыдущей задачи будем предполагать, что после отказа элемент восстанавливается в течение случайного времени η_i. За это время отказ обнаруживается, и элемент либо ремонтируется, либо заменяется новым. Таким образом рассматривается случайный процесс (рис 14.1), образованный временами жизни τ_i и временами восстановления η_i. Закон распределения всех времен τi одинаков, также одинаков закон распределения всех времен ηi. Величины τi и ηi независимы.



Процесс конечным временем восстановления
Рис. 14.1

По заданным законам распределения определяются среднее время работы элемента $\mathbf{T_0 = M[\tau]}$ и среднее время ремонта $\mathbf{T_p = M[\eta]}$.

Для независимых τ и η среднее время между отказами равно $\mathbf{T = T_0 + T_p}$.

Не повторяя вывода, разобранного в предыдущем разделе, сразу укажем, что число отказов $\nu(\mathbf{t})$ за время $\mathbf{t}$ при больших значениях времени распределено асимптотически нормально со средним $\dfrac{\mathbf{t}}{\mathbf{T_0 + T_p}}$ и дисперсией

$\dfrac{(\sigma_0^2 + \sigma_p^2)\mathbf{t}}{(\mathbf{T_0 + T_p})^3}$, где σ_0^2 и σ_p^2 - дисперсии величин τ и η соответственно.

Важной характеристикой для данного процесса вероятность исправного состояния элемента. Эта вероятность легко вычисляется, если закон надежности элемента экспоненциальный

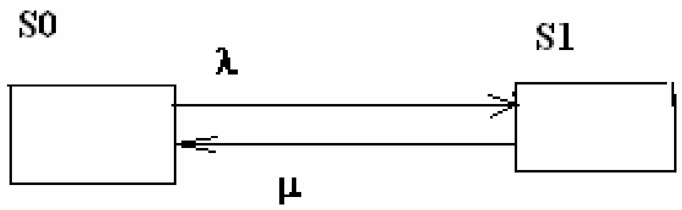
$$\mathbf{P_0(t) = e^{-\lambda t}}$$

и интегральный закон распределения времени восстановления также экспоненциальный

$$Q_p(t) = e^{-\mu t}$$

Принятый закон восстановления является некоторой абстрактной моделью. Экспоненциальный закон надежности предполагает модель мгновенных повреждений, то есть время отказа есть независимая от предыстории случайная величина. Следовательно, время ремонта, поскольку принят экспоненциальный закон, также является величиной, независимой от того, как и сколько времени проводился ремонт. Здесь можно принять следующую модель. Ремонт проводится путем случайного выбора очередной части для проверки элемента. Проверяющий как бы не знает ни устройства, ни логики работы элемента. Приблизительность модели окупается выгодами применения экспоненциального закона.

Граф процесса при конечном времени восстановления показан на рис 14.2.



Граф системы с конечным временем восстановления
Рис. 14.2

В состоянии $\mathbf{S0}$ элемент исправен. Когда происходит отказ, элемент переходит в состояние $\mathbf{S1}$. В этом состоянии элемент находится, пока не оканчивается ремонт. Тогда он снова переходит в состояние $\mathbf{S0}$ и т. д. Условные плотности вероятности переходов при экспоненциальных законах отказов и восстановлений равны, как это было показано в разделе 12, опасности отказа λ и условной плотности μ. Процесс смены состояний образует непрерывную цепь Маркова. По графу процесса составляем уравнения Колмогорова

$$\begin{aligned}\dot{\mathbf{P}}_0 &= -\lambda \mathbf{P}_0 + \mu \mathbf{P}_1 \\ \dot{\mathbf{P}}_1 &= \lambda \mathbf{P}_0 - \mu \mathbf{P}_1\end{aligned} \qquad (14.1)$$

Здесь $\mathbf{P}_0$ - вероятность исправного состояния, которую и необходимо найти. Воспользуемся для понижения порядка системы условием

$$\mathbf{P}_0 + \mathbf{P}_1 = 1$$

Подставляем в первое уравнение (14.1) это условие, получим

$$\dot{\mathbf{P}}_0 = -\lambda \mathbf{P}_0 + \mu(1 - \mathbf{P}_0)$$

или

$$\dot{\mathbf{P}}_0 + (\lambda + \mu)\mathbf{P}_0 = \mu$$

Это уравнение надо решить при начальном условии

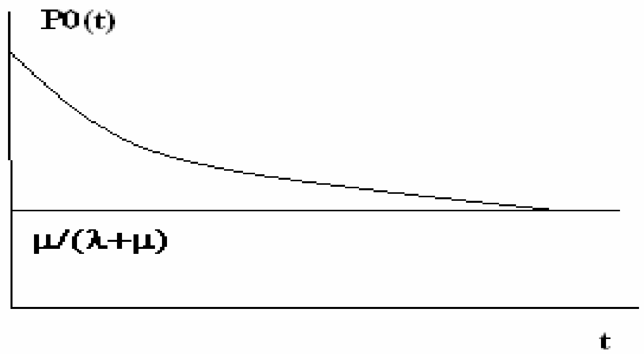
$$\mathbf{P}_0(0) = 1$$

Частное решение уравнения равно

$$\mathbf{P}_0{}^* = \frac{\mu}{\lambda + \mu} \qquad (14.2)$$

Общее решение или вероятность исправного состояния равно

$$\mathbf{P}_0(\mathbf{t}) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}\mathbf{e}^{-(\lambda + \mu)\mathbf{t}} \qquad (14.3)$$



Зависимость вероятности исправного состояния от времени

Рис.14.3

Зависимость вероятности исправного состояния от времени показана на рис. 14.3. В начальный момент элемент исправен – вероятность равна единице.

Затем вероятность исправного состояния уменьшается и в пределе стремится к постоянному установившемуся значению

$$\mathbf{k_г = P_0(\infty)} = \frac{\mu}{\lambda + \mu} \qquad (14.4)$$

которое называется коэффициентом готовности.

Так как среднее время жизни и среднее время ремонта при экспоненциальном законе

$$\mathbf{T_0} = \frac{\mathbf{1}}{\lambda}$$

$$\mathbf{T_p} = \frac{\mathbf{1}}{\mu}$$

то коэффициент готовности равен отношению среднего времени функционирования элемента к общему времени эксплуатации:

$$\mathbf{k_г} = \frac{\mathbf{T_0}}{\mathbf{T_0 + T_p}}. \qquad (14.5)$$

Выражения (14.4) или (14.5) характеризуют такой режим, когда в процессе отказов и восстановлений элемента, то есть в процессе смены состояний, вероятности состояний принимают установившиеся значения.

Если в организации процесса эксплуатации необходимо обеспечить, чтобы $\mathbf{N}$ элементов (автобусов, ракет, автоматов по продаже газет) в среднем было работоспособно, общее число элементов $\mathbf{N_0}$, запущенных в эксплуатацию, должно подсчитываться из условия. Оценка вероятности исправного состояния равна

$$\mathbf{P_0(\infty) = k_г} \approx \frac{\mathbf{N}}{\mathbf{N_0}}$$

Тогда

$$\mathbf{N_0} = \frac{\mathbf{N}}{\mathbf{k_г}} = \mathbf{N}\frac{\mathbf{T_0 + T_p}}{\mathbf{T_0}} \qquad (14.6)$$

Очевидно, что чем меньше среднее время ремонта, тем меньше элементов надо закрепить за рассматриваемым процессом эксплуатации.

Контрольные вопросы

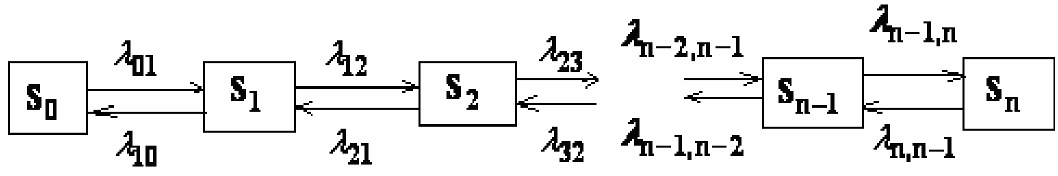
1. Опишите принятую в этом разделе модель восстановления, её достоинства и недостатки.

2. Что такое коэффициент готовности?

3. Какова достоверность достаточности количества элементов для организации процесса эксплуатации, рекомендованная в этом разделе?

15. Резервированная система с конечным временем восстановления

Пусть система состоит из $\mathbf{n}$ элементов: одного основного блока и $\mathbf{n-1}$ элементов, работающих в горячем резерве. Всего имеется $\mathbf{n}$ элементов и пока хотя бы один исправен, система работает. Отказавшие блоки восстанавливаются. Процесс отказов и ремонта опишем позднее.



Граф резервированной системы с восстановлением
Рис. 15.1.

Предполагаем, что система марковская. Состояния системы обозначим через $\mathbf{S_i}$, где $\mathbf{i}$ - число неисправных блоков. Граф системы показан на рис. 15.1. Очевидно, что в процессе работы элементы отказывают, восстанавливаются, причем в ремонте одновременно могут находиться несколько элементов, в том числе и резервные, так как они находятся в горячем резерве, где отказ возможен. То есть система переходит из одного состояния в другое. Условные плотности вероятности отказов и окончания ремонта пока запишем в общем виде λ_{ij}, которая соответствует переходу из состояния $\mathbf{S_i}$ в состояние $\mathbf{S_j}$. Граф, изображенный на рис. 15.1, называется «схемой гибели и размножения», используемой во многих отраслях.

Составим уравнения Колмогорова для этой системы

$$\dot{P}_0 = -\lambda_{01}P_0 + \lambda_{10}P_1$$
$$\dot{P}_1 = \lambda_{01}P_0 - (\lambda_{12} + \lambda_{10})P_1 + \lambda_{21}P_2 ,$$

и т. д.

Наиболее часто в этих системах интерес представляют установившиеся значения вероятностей состояний

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

которые определяются сравнительно несложно. В установившемся состоянии параметры (вероятности) не меняются и их производные равны нулю. Следовательно, установившиеся значения вероятностей определяются из уравнений Колмогорова, в которых значения производных приняты нулевыми. Система дифференциальных уравнений превращается в алгебраическую систему уравнений. Установившиеся значения здесь будут существовать, если граф системы содержит конечное число состояний и из каждого состояния за конечное число шагов можно перейти в любое другое состояние.

Запишем алгебраические уравнения для установившихся вероятностей. Для состояния S_0 имеем

$$\lambda_{01}P_0 = \lambda_{10}P_1 \tag{15.1}$$

Для состояния S_1 получим

$$\lambda_{10}P_1 + \lambda_{12}P_1 = \lambda_{01}P_0 + \lambda_{21}P_2 \tag{15.2}$$

Получено равенство сумм, соответствующих входящим и выходящим из состояния S_1 стрелкам. Вычитая из обеих частей (15.2), значения из (15.1). получим

$$\lambda_{12}P_1 = \lambda_{21}P_2 \tag{15.3}$$

Для состояния S_2 имеем

$$\lambda_{21}P_2 + \lambda_{23}P_2 = \lambda_{12}P_1 + \lambda_{32}P_3$$

Вычитая из обеих частей (15.3), получим

$$\lambda_{23}P_2 = \lambda_{32}P_3 \tag{15.4}$$

Видно, что алгебраические уравнения (15.1), (15.3), (15.4) связывают выражения, соответствующие парам стрелок, стоящим друг под другом. Поэтому для состояний сразу можем записать

$$\lambda_{01}P_0 = \lambda_{10}P_1$$

$$\lambda_{12}P_1 = \lambda_{21}P_2$$

$$\lambda_{23}P_2 = \lambda_{32}P_3$$

………………..

$$\lambda_{n-1,n}P_{n-1} = \lambda_{n,n-1}P_n$$

Решаем эту систему последовательно, исключая переменные

$$P_1 = \frac{\lambda_{01}}{\lambda_{10}}P_0$$

$$P_2 = \frac{\lambda_{12}}{\lambda_{21}}P_1 = \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}}P_0 \tag{15.5}$$

$$P_3 = \frac{\lambda_{23}}{\lambda_{32}}P_2 = \frac{\lambda_{23}\lambda_{12}\lambda_{01}}{\lambda_{32}\lambda_{21}\lambda_{10}}P_0$$

Здесь также ясна структура выражений. Для вероятности каждого состояния в числителе стоит произведение условных плотностей вероятности переходов от начала графа в данное состояние, а в знаменателе – произведение условных плотностей вероятности от данного состояния в начало графа. Поэтому сразу можем для вероятности состояния S_n записать

$$P_n = \frac{\lambda_{01}\lambda_{12}\lambda_{23}\ldots\lambda_{n-2,n-1}\lambda_{n-1,n}}{\lambda_{n,n-1}\lambda_{n-1,n-2}\ldots\lambda_{32}\lambda_{21}\lambda_{10}}P_0 \tag{15.6}$$

Заметим, что вероятность последнего состояния – это вероятность отказа системы, так как все элементы системы неисправны и находятся в ремонте.

Таким образом, вероятности всех состояний выражены через вероятность начального состояния S_0.

Запишем условие

$$P_0 + P_1 + P_2 + \ldots + P_n = 1$$

Подставим в него выражения (15.5), (15.6)

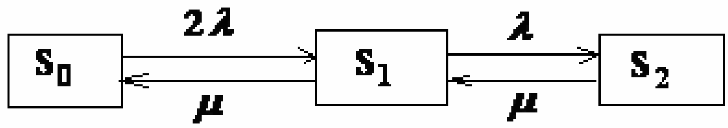
$$\mathbf{P_0} + \frac{\lambda_{01}}{\lambda_{10}}\mathbf{P_0} + \frac{\lambda_{01}\lambda_{12}}{\lambda_{21}\lambda_{10}}\mathbf{P_0} + \ldots \frac{\lambda_{01}\lambda_{12}\ldots\lambda_{n-1,n}}{\lambda_{n,n-1}\ldots\lambda_{21}\lambda_{10}}\mathbf{P} = 1$$

Отсюда установившееся значение начального состояния $\mathbf{P_0}$ равно

$$\mathbf{P_0} = \frac{1}{1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{01}\lambda_{12}}{\lambda_{21}\lambda_{10}} + \frac{\lambda_{01}\lambda_{12}\lambda_{23}}{\lambda_{32}\lambda_{21}\lambda_{10}} + \ldots + \frac{\lambda_{01}\lambda_{12}\lambda_{23}\ldots\lambda_{n-1,n}}{\lambda_{n,n-1}\ldots\lambda_{32}\lambda_{21}\lambda_{10}}} \qquad (15.7)$$

Все остальные установившиеся вероятности состояний выражаются через $\mathbf{P_0}$ с помощью соотношений (15.5),(15.6). Получено в общем виде решение для установившихся значений в «схеме гибели и размножения».

Перейдем к решению задачи, поставленной в начале раздела. Пусть имеется один основной элемент и один элемент в «горячем» резерве. При отказе одного из них отказ мгновенно обнаруживается и начинается восстановление Закон надежности и закон восстановления – экспоненциальные. Ремонтный участок – один. Интенсивность восстановления на этом участке равна μ. Поэтому, если отказали оба элемента, восстанавливается только один элемент, другой ожидает своей очереди, то есть интенсивность восстановления остается равной μ. Граф этой системы показан на рис. 15.2.



Система восстановления с одним резервным элементом

Рис. 15.2

Найдем установившиеся значения вероятностей состояний:

$\mathbf{S_0}$ - все элементы исправны.

$\mathbf{S_1}$ - один элемент отказал и восстанавливается, второй элемент работает.

$\mathbf{S_2}$ - оба элемента отказали, один элемент продолжает восстанавливаться, второй элемент ждет очереди на восстановление. Система неработоспособна.

Условная плотность вероятности перехода $\lambda_{01} = 2\lambda$, так как этот переход происходит при отказе одного из двух исправных элементов. Сравнивая обозначения на рис 15.1 и рис 15.5, из (15.5) – (15.7) получаем

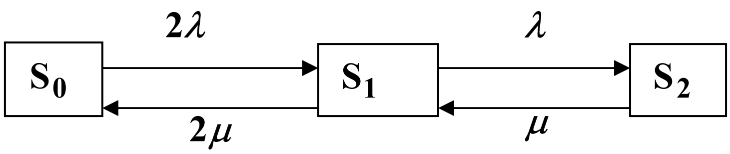
$$\mathbf{P_0} = \frac{1}{1 + \frac{2\lambda}{\mu} + \frac{2\lambda^2}{\mu^2}} = \frac{\mu^2}{\mu^2 + 2\lambda\mu + 2\lambda^2} \qquad (15.8)$$

$$\mathbf{P_1} = \frac{2\lambda}{\mu}\mathbf{P_0} = \frac{2\lambda\mu}{\mu^2 + 2\lambda\mu + 2\lambda^2} \qquad (15.9)$$

Коэффициент готовности – вероятность исправного состояния

$$\mathbf{k_{г1}} = \mathbf{P_0} + \mathbf{P_1} = \frac{2\lambda\mu + \mu^2}{\mu^2 + 2\lambda\mu + 2\lambda^2} \qquad (15.10)$$

Рассмотрим теперь процесс, когда в системе есть два ремонтных участка и в том случае, когда отказали оба элемента, каждый из них сразу, без ожидания ремонтируется на своем участке. Граф состояний показан на рис.15.3. Условная плотность вероятности перехода $\lambda_{21} = 2\mu$, так как при этом восстанавливается один из двух ремонтируемых элементов.



Граф системы с двумя ремонтными участками

Рис. 15.3

Способ решения в этой и предыдущей задаче совпадают. Для начального состояния получим

$$\mathbf{P_0} = \frac{1}{1 + \frac{2\lambda}{\mu} + \frac{\lambda^2}{\mu^2}} = \frac{\mu^2}{(\lambda + \mu)^2} \qquad (15.11)$$

Установившееся значение вероятности для состояния $\mathbf{S_1}$ равно

$$\mathbf{P_1} = \frac{2\lambda}{\mu}\mathbf{P_0} = \frac{2\lambda\mu}{(\lambda + \mu)^2} \qquad (15.12)$$

Коэффициент готовности равен

$$\mathbf{k_{г2}} = \mathbf{P_0} + \mathbf{P_1} = \frac{2\lambda\mu + \mu^2}{(\lambda + \mu)^2} \qquad (15.13)$$

Обозначим $\rho = \frac{\lambda}{\mu}$ и сравним все три рассмотренных случая систем с конечным временем восстановления. При $\rho << 1$ (интенсивность ремонта много больше интенсивности отказов) имеем

Нерезервированная система

$$\mathbf{k_{г}} = \frac{\mu}{\lambda + \mu} = \frac{1}{1 + \rho} \approx 1 - \rho \qquad (15.14)$$

Резервированная система с одним ремонтным участком

$$\mathbf{k_{г1}} = \frac{1 + 2\rho}{1 + 2\rho + 2\rho^2} \approx 1 - \frac{2\rho^2}{1 + 2\rho} \approx 1 - 2\rho^2 \qquad (15.15)$$

Резервированная система с двумя ремонтными участками

$$\mathbf{k_{г2}} = \frac{1 + 2\rho}{1 + 2\rho + \rho^2} \approx 1 - \frac{\rho^2}{1 + 2\rho} \approx 1 - \rho^2 \qquad (15.16)$$

Очевидны преимущества, которые дают как резервирование, так и повышение интенсивности восстановления.

Контрольные вопросы

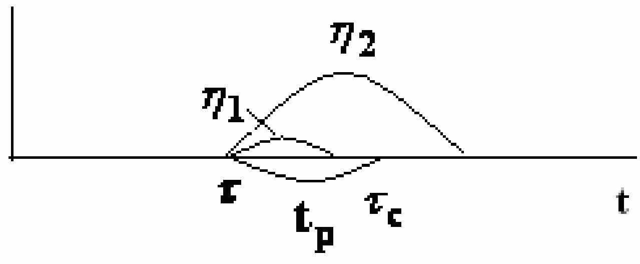
1. Нарисуйте граф «схемы гибели и размножения», объясните его структуру.

2. Запишите формулы вычисления установившихся значений вероятностей состояния «схемы гибели и размножения».

3. Нарисуйте граф состояний резервированных систем с одним и двумя участками восстановления.

16. Системы с резервом времени

Напоминаем, что системами с резервом времени называются системы, которые имеют некий резерв, за счет которого они поддерживают работоспособное состояние системы во время восстановительного ремонта. Этот резерв ограничен во времени. Процедура восстановления систем с резервом времени поясняется на рис. 16.1.



Отказ и восстановление систем с резервом времени

Рис. 16.1

τ - момент отказа системы,

$\mathbf{t_p}$ - величина резерва времени,

$\eta_1 < \mathbf{t_p}$ - время ремонта меньше резерва времени, отказ системы не наступил,

$\eta_2 > \mathbf{t_p}$ - время ремонта больше резерва времени, отказ системы наступил в момент времени τ_c.

Принимаем законы отказов и восстановлений экспоненциальными. Интенсивность ремонта μ постоянна и не зависит от времени окончания запаса.

Интегральный закон распределения времени восстановления при экспоненциальном законе восстановления имеет вид

$$\mathbf{P_в(t) = P[\eta < t] = 1 - e^{-\mu t}}$$

Дифференциальный закон распределения η - плотность вероятности равна

$$\mathbf{w}(\eta)=\frac{\mathbf{dQ_в(t)}}{\mathbf{dt}}=\mu\mathbf{e}^{-\mu\mathbf{t}} \tag{16.1}$$

Время неработоспособности равно $\eta-\mathbf{t_p}$, если $\eta>\mathbf{t_p}$ и время неработоспособности равно нулю, если $\eta<\mathbf{t_p}$. Тогда среднее время неработоспособности равно

$$\mathbf{T_{ot}}=\int_{\mathbf{t_p}}^{\infty}(\eta-\mathbf{t_p})\mathbf{w}(\eta)\mathbf{d}\eta=\int_{\mathbf{t_p}}^{\infty}(\eta-\mathbf{t_p})\mu\mathbf{e}^{-\mu\eta}\mathbf{d}\eta$$

Проведем замену переменных

$$\eta-\mathbf{t_p}=\mathbf{x}$$

Тогда нижний предел при $\eta=\mathbf{t_p}$ равен нулю. Верхний предел равен бесконечности. Дифференциал новой переменной интегрирования равен $\mathbf{dx}=\mathbf{d}\eta$.

Получаем

$$\mathbf{T_{ot}}=\int_{\mathbf{0}}^{\infty}\mathbf{x}\mu\mathbf{e}^{-\mu(\mathbf{x}+\mathbf{t_p})}\mathbf{dx}=\mu\mathbf{e}^{-\mu\mathbf{t_p}}\int_{\mathbf{0}}^{\infty}\mathbf{x}\mathbf{e}^{-\mu\mathbf{x}}\mathbf{dx}$$

Интегрируя интеграл по частям, получаем выражение для среднего времени неработоспособности

$$\mathbf{T_{ot}}=\frac{\mathbf{1}}{\mu}\mathbf{e}^{-\mu\mathbf{t_p}} \tag{16.2}$$

За один цикл процесса примем время работы элемента до отказа плюс время ремонта. За время ремонта элемент может быть работоспособен все это время или частично. Среднее время неработоспособности вычислено.

Среднее время цикла равно

$$\mathbf{T_ц}=\mathbf{T_0}+\mathbf{T_p}=\frac{\mathbf{1}}{\lambda}+\frac{\mathbf{1}}{\mu} \tag{16.3}$$

Коэффициент готовности как отношение среднего времени работоспособного состояния к среднему времени цикла равно

$$\mathbf{k_г} = \frac{\mathbf{T_ц - T_{ot}}}{\mathbf{T_ц}} = \frac{\frac{1}{\lambda} + \frac{1}{\mu} - \frac{\mathbf{e}^{-\mu \mathbf{t_p}}}{\mu}}{\frac{1}{\lambda} + \frac{1}{\mu}}$$

Используем ранее введенное понятие $\rho = \frac{\lambda}{\mu}$. Получим для коэффициента готовности выражение

$$\mathbf{k_г} = \frac{1 + \rho(1 - \mathbf{e}^{-\mathbf{k}})}{1 + \rho} \qquad (16.4)$$

где $\mathbf{k} = \mathbf{t_p}\mu = \frac{\mathbf{t_p}}{\mathbf{T_p}}$ - отношение резерва времени к среднему времени ремонта.

При $\mathbf{k} = 0$ (нет резерва времени) получаем известный результат для нерезервированных систем

$$\mathbf{k_г} = \frac{1}{1 + \rho}$$

При $\mathbf{k} >> 1$ (резерв времени практически всегда обеспечивает работоспособность системы во время ремонта) получаем очевидный результат

$$\mathbf{k_г} \to 1$$

Система всегда работоспособна.

Рассмотрим, какими параметрами должна обладать системе с резервом времени, чтобы быть адекватной системе с обычным резервированием. Разложим в ряд (16.4) по малым значениям ρ. С точностью до членов второго порядка малости получим

$$\mathbf{k_г} = 1 - \rho(1 - \rho)\mathbf{e}^{-\mathbf{k}}$$

Приравнивая полученный результат коэффициенту готовности системы с обычным резервированием (один ремонтный участок), получим

$$1 - \rho(1 - \rho)\mathbf{e}^{-\mathbf{k}} \approx 1 - \rho^2,$$

или $\mathbf{e^{-k}} \approx \rho$. Следовательно, отношение резерва по времени к среднему времени ремонта должно быть равно

$$\mathbf{k = \ln \frac{1}{\rho}}$$

Для оценки примем $\rho = \mathbf{10^{-2}}$. Получим $\mathbf{k = 4.6}$. Во столько раз величина резерва по времени должна быть больше среднего времени ремонта.

Контрольные вопросы

1. Что такое резервирование по времени? Приведите примеры.

2. В каком случае система остается работоспособной в течение всего времени ремонта?

3. Напишите выражение для коэффициента готовности системы с резервированием по времени.

17. Основные понятия систем массового обслуживания (СМО). Одноканальная СМО с отказами

Примерами систем массового обслуживания (СМО) могут служить такие системы по обслуживанию, как телефонные станции, магазины, системы поражения воздушных целей, бензоколонки и т. д. Характеристиками СМО являются:

Каналы обслуживания. Каналы обслуживания это обслуживающие устройства, такие как каналы связи, продавцы и кассиры, системы слежения и поражения воздушных целей, раздаточные бензоколонки и т. д.

Входной поток заявок. Каждая СМО предназначена для обслуживания какого-то потока заявок – последовательности однородных событий, следующих одно за другим в моменты времени с известными статистическими характеристиками, такие как запросы на телефонные соединения, появление покупателей, появление воздушных целей, заезд на заправку автомобилей и т. д.

Правила обслуживания. Под правилами обслуживания будем понимать правила обслуживания заявки:

В порядке очереди, первый пришел – первым обслуживаешься,

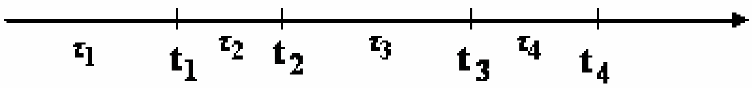
Есть приоритетные заявки, обслуживаемые без очереди,

Стековая очередь, первым пришел – последним обслуживаешься.

Организация очереди. В СМО могут допускаться очереди на обслуживание, когда заявка пришла, но канал обслуживания занят и заявке разрешается ждать в системе, пока не найдется канал, который будет заявку обслуживать. Для этого в технической системе предусматривают специальные накопительные устройства , где заявки ожидают своей очереди. Тогда очередь будет ограниченной длины, в ней ограниченное число мест. Если очередная заявка придет, а все места в очереди заняты, заявка покинет систему, будет потеряна для обслуживания. Есть системы, где допускаются очереди неограниченной длины. В таких системах потерь заявок нет. Есть системы, где очереди не допускаются.

Организация обслуживания. Здесь рассматриваются такие характеристики, как число каналов обслуживания, вероятностный закон обслуживания заявки – распределение времени обслуживания.

Рассмотрим здесь несколько подробнее такое понятие как потоки заявок. На Рис 17.1 показан пример потока заявок. Заявки поступают в моменты времени t_i, интервал времени между заявками равен τ_i. Поток характеризуется либо законом распределения $\mathbf{t}$, либо законом распределения τ.



Поток событий

Рис. 17.1

Поток *регулярный*, если интервалы между заявками одинаковы $\tau_i = \mathbf{Const}$.

Поток *случайный*, если интервалы между заявками непрерывные случайные величины.

Поток *стационарный*, если статистические свойства потока не зависят от времени.

Поток *без последействия*, если свойства потока в текущий момент времени не зависят от его предыдущей истории.

Поток *ординарный*, если вероятность прихода двух и более заявок в один и тот же момент времени бесконечно мала.

Случайный поток, стационарный, без последействия и ординарный есть *простейший* или *пуассоновский* поток. О нем несколько раз уже упоминалось в курсе. Вероятность прихода $\mathbf{k}$ заявок за время $\mathbf{T}$ для пуассоновского потока равна

$$\mathbf{P_k(T)} = \frac{(\alpha \mathbf{T})^{\mathbf{k}}}{\mathbf{k!}} \mathbf{e}^{-\alpha \mathbf{T}} \qquad (17.1)$$

Здесь α - средняя частота наступления событий.

Упоминалось, что модель мгновенных повреждений при экспоненциальном законе распределения отказов и определение пуассоновского потока идентичны. Процесс отказов с мгновенным восстановлением это пуассоновский поток событий. Поэтому

$$\alpha = \lambda \tag{17.2}$$

Преимущество применения пуассоновского потока заключается в том, что при сложении большого числа стационарных и ординарных потоков образуется пуассоновский поток.

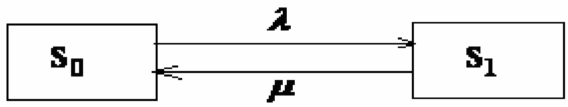
Потоком Эрланга *n* – го порядка называется поток, образованный каждой *n* – ой заявкой пуассоновского потока. Если **n** велико, поток превращается в нормальный (см. Г – распределение времени безотказной работы).

Потоком Пальма называется поток, в котором промежутки времени между заявками представляют собой независимые, одинаково распределенные случайные величины, например, корабли, идущие кильватерной колонной.

Рассмотрим простейшую задачу из теории СМО – задачу об одноканальной СМО с отказами. В системе имеется один канал обслуживания. Входной поток пуассоновский с интенсивностью $\lambda = \lambda(\mathbf{t})$. Очереди в системе не допускаются и заявка, заставшая канал занятым, покидает систему необслуженной. Канал обслуживает заявку случайное время τ. Принимаем, что эта случайная величина распределена по экспоненциальному закону

$$\mathbf{q}(\tau) = \mu \mathbf{e}^{-\mu\tau} \tag{17.3}$$

На Рис. 17.2 показан граф состояний системы. Состояние $\mathbf{S_0}$- канал обслуживания свободен, в системе нет заявок, она простаивает. Состояние $\mathbf{S_1}$ - в систему поступила заявка, канал её обслуживает, он занят. Если в систему, когда она находится в этом состоянии, поступит заявка, она застанет канал занятым, очередь в системе запрещена, ожидать заявке негде. Заявка покидает систему необслуженной, для системы заявка потеряна. Поэтому вероятность состояния $\mathbf{S_1}$ - это вероятность потери заявки.



Граф одноканальной СМО с отказами
Рис. 17.2

Составим уравнения Колмогорова

$$\begin{aligned} \dot{P}_0 &= -\lambda P_0 + \mu P_1 \\ \dot{P}_1 &= -\mu P_1 + \lambda P_0 \end{aligned} \tag{17.4}$$

Используем соотношение

$$P_0 + P_1 = 1$$

Получим вместо (17.4) дифференциальное уравнение первого порядка

$$\dot{P}_0 = -(\lambda + \mu)P_0 + \mu \tag{17.5}$$

Решение этого уравнения при начальных условиях

$$\begin{aligned} P_0(0) &= 1 \\ P_1(0) &= 0 \end{aligned}$$

равно

$$P_0 = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t} \tag{17.6}$$

Вероятность P_0 - это вероятность обслуживания заявки, так как в этом состоянии канал свободен, и пришедшая заявка поступает на обслуживание. Вероятность обслуживания заявки называется относительной пропускной способностью q. Такая часть заявок обслуживается в среднем.

Рассматривая установившиеся значения, получим

$$q = \frac{\mu}{\lambda + \mu} \tag{17.7}$$

Абсолютной пропускной способностью канала **A** называется количество заявок, которое он обслуживает в единицу времени. В единицу времени приходит всего λ заявок, а обслуживается

$$\mathbf{A} = \lambda \mathbf{q} \tag{17.8}$$

Вероятность потери заявки равна

$$\mathbf{Рппо} = \mathbf{1} - \mathbf{q}$$

Для установившегося значения вероятности потери заявки получим

$$\mathbf{P_{пот}} = \frac{\lambda}{\lambda + \mu} \tag{17.9}$$

Контрольные вопросы

1. Чем определяется система массового обслуживания?

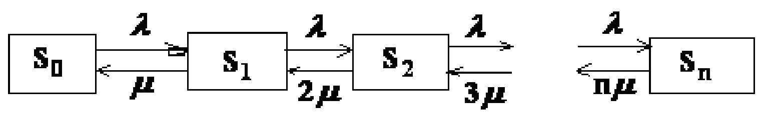
2. Что такое пуассоновский поток событий?

3. Нарисуйте граф одноканальной СМО с отказами. Дайте характеристику состояниям системы.

4. Чему равны относительная и абсолютная пропускные способности одноканальной СМО с отказами?

18. Многоканальная СМО с отказами

Пусть система содержит $\mathbf{n}$ каналов обслуживания (продавцов, бензоколонок, каналов связи).Интенсивности обслуживания каждым каналом одинаковы и равны μ. Очереди в системе не допускаются. Если в системе находятся $\mathbf{m}$ заявок, причем $\mathbf{m} \le \mathbf{n}$, то они обслуживаются $\mathbf{m}$ каналами с интенсивностью $\mathbf{m}\mu$. Обозначим состояния системы $\mathbf{S_i}$, где $\mathbf{i}$ - число заявок в системе и соответственно число занятых каналов. Граф состояний системы показан на рис 18.1. Если система находится в состоянии $\mathbf{S_n}$ (все каналы обслуживания заняты), то очередные поступающие заявки не могут быть обслужены, очереди не допускаются, и они покидают систему необслуженными, теряются.



Многоканальная СМО с отказами
Рис. 18.1

Для установившегося значения вероятности $\mathbf{P_0}$, используя результаты, полученные для «схемы гибели и размножения», имеем

$$\mathbf{P_0} = \frac{1}{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{2 * 3\mu^3} + ... + \frac{\lambda^n}{n!\,\mu^n}} \qquad (18.1)$$

Используя ранее введенное обозначение

$$\rho = \frac{\lambda}{\mu}, \qquad (18.2)$$

Получим выражение для установившегося значения вероятности простоя канала связи в виде

$$\mathbf{P_0} = [1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + ... + \frac{\rho^n}{n!}] \qquad (18.3)$$

Вероятность потери заявок равна

$$\mathbf{P_n} = \frac{\rho^{\mathbf{n}}}{\mathbf{n!}}\mathbf{P_0} \tag{18.4}$$

Очевидно, что вероятность быть обслуженным или относительная пропускная способность равна

$$\mathbf{q = 1 - P_n} \tag{18.5}$$

Абсолютная пропускная способность равна

$$\mathbf{A} = \lambda\mathbf{q} = \lambda(\mathbf{1 - P_n}) \tag{18.6}$$

Среднее число занятых каналов $\mathbf{\bar{k}}$ можно определить следующим образом. За единицу времени система обслуживает в среднем $\mathbf{A}$ заявок. Один канал за единицу времени обслуживает в среднем μ заявок. Следовательно,

$$\mathbf{\bar{k}} = \frac{\mathbf{A}}{\mu} = \frac{\lambda}{\mu}(\mathbf{1 - P_n}) = \rho(\mathbf{1 - P_n}) \tag{18.7}$$

Полученные соотношения позволяют выбрать количество оборудования $\mathbf{n}$ и качество оборудования μ, исходя из различных соотношений между вероятностью не справиться с обслуживанием (потерять доход от потери заявки) и вероятностью простоя оборудования.

Поставим следующую задачу. За каждую необслуженную заявку выплачивается штраф в размере $\mathbf{a}$ руб. (??? Изменить постановку задачи. За каждую обслуженную заявку получают доход???). При простое одного канала убыток составляет $\mathbf{c}$ руб. Необходимо выбрать число каналов, чтобы средняя величина потерь была минимальна.

Вероятность отказа при обслуживании равна (18.4). Среднее число необслуженных заявок в единицу времени равно $\lambda\mathbf{P_n}$ и средний размер штрафа в единицу времени равен

$$\mathbf{S_1} = \mathbf{a}\lambda\mathbf{P_n}$$

Среднее число незагруженных каналов равно $\mathbf{n - \bar{k}}$. Тогда средняя величина убытка от простоя каналов в единицу времени равна

$$\mathbf{S_2 = c(n - \bar{k}) = c[n - \rho(1 - P_n)]}$$

Условие минимума потерь запишем в виде

$$\mathbf{S = S_1 + S_2 = a\lambda P_n + c[n - (1 - P_n)] = min}$$

В таблице 18.1 приведены результаты расчетов зависимости потерь от числа каналов. Расчеты проведены для произвольно выбранных значений:

c=1, a=10, $\boldsymbol{\rho = 1}$

Таблица 18.1

n	$\mathbf{P_n}$	$\mathbf{\bar{k}}$	**S**
1	**0.5**	**0.5**	**5.5**
2	**0.2**	**0.8**	**3.2**
3	**0.062**	**0.938**	**2.7**
4	**0.002**	**0.998**	**3.0**

Видно, что число каналов следует выбрать равным трем или четырем. Обратите внимание, что с увеличением числа каналов среднее число занятых каналов приближается к единице. Действительно, при $\lambda = \mu$ мощности одного канала достаточно, чтобы справиться с обслуживанием потока заявок той же мощности. Но если мы выберем один канал, половина заявок останется необслуженными, а канал половину времени будет простаивать. С увеличением числа каналов увеличивается вероятность обслуживания за счет увеличения суммарного времени простоя каналов (увеличивается вероятность для заявки застать какой – либо канал свободным).

Контрольные вопросы

1. Нарисуйте граф многоканальной СМО с отказами. Поясните характеристики графа.

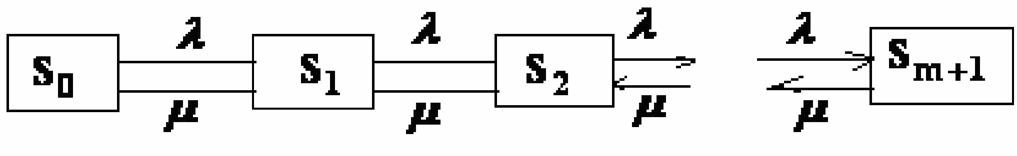
2. Что такое среднее число занятых каналов?

3. Чему равны относительная и абсолютная пропускные способности одноканальной СМО с отказами?

19. Одноканальная СМО с ожиданием

Пусть система имеет один канал с интенсивностью обслуживания μ. Заявка, поступившая в момент, когда канал занят, становится в очередь и ожидает обслуживания. Предположим вначале, что количество мест в очереди ограничено и равно $\mathbf{m}$. То есть заявка, поступившая в систему, когда в очереди стоит $\mathbf{m}$ заявок (все места в очереди заняты), покидает систему необслуженной, теряется. Такая ситуация может возникнуть при использовании буферного запоминающего устройства (БЗУ) с числом ячеек $\mathbf{m}$. При переполнении БЗУ посту пающие отсчеты теряются и в дальнейшей обработке информации не участвуют.

Обозначим состояния системы через $\mathbf{S_i}$, где $\mathbf{i}$ - число заявок, находящихся в системе. Граф состояний системы показан на Рис. 19.1.



Граф одноканальной СМО с ожиданием
Рис. 19.1

Состояние $\mathbf{S_0}$ - в системе нет заявок, канал свободен и простаивает.

Состояние $\mathbf{S_1}$ - канал занят, очереди нет.

Состояние $\mathbf{S_2}$ - канал занят, одна заявка обслуживается, одна заявка в очереди.

……..

Состояние $\mathbf{S_{m+1}}$ - канал занят, одна заявка обслуживается, $\mathbf{m}$ заявок в очереди. Очередные заявки, поступившие в систему, когда система находится в этом состоянии, не обслуживаются, теряются.

Все заявки обслуживаются одним каналом, поэтому интенсивность восстановления всегда одинакова и равна μ.

Согласно соотношениям, полученным в «схеме гибели и размножения», установившееся значение вероятности простоя канала равно

$$\mathbf{P_0} = \frac{\mathbf{1}}{\mathbf{1} + \frac{\lambda}{\mu} + \frac{\lambda^{\mathbf{2}}}{\mu^{\mathbf{2}}} + \mathbf{....} + \frac{\lambda^{\mathbf{m+1}}}{\mu^{\mathbf{m+1}}}} \qquad (19.1)$$

Или, используя (18.2)

$$\mathbf{P_0} = \frac{\mathbf{1}}{\mathbf{1} + \rho + \rho^{\mathbf{2}} + \mathbf{...} + \rho^{\mathbf{m+1}}} \qquad (19.2)$$

Выражение в знаменателе есть геометрическая прогрессия. Сумма $\mathbf{m+2}$ её членов равна

$$\sum_{\mathbf{m+2}} = \frac{\mathbf{a_1(s^{m+2} - 1)}}{\mathbf{s - 1}}$$

Где

$\mathbf{a_1 = 1}$ - первый член прогрессии,

$\mathbf{S} = \rho$ - множитель прогрессии.

Тогда

$$\sum_{\mathbf{m+2}} = \frac{\rho^{\mathbf{m+2}} - \mathbf{1}}{\rho - \mathbf{1}}$$

Следовательно, вероятность простоя канала равна

$$\mathbf{P_0} = \frac{\mathbf{1} - \rho}{\mathbf{1} - \rho^{\mathbf{m+2}}} \qquad (19.3)$$

Остальные вероятности, согласно (???) равны

$$\mathbf{P_1} = \rho \mathbf{P_0}$$

$$\mathbf{P_1} = \rho^{\mathbf{2}} \mathbf{P_0} \qquad (19.4)$$

……………

$$\mathbf{P_{m+1}} = \rho^{\mathbf{m+1}} \mathbf{P_0}$$

Важнейшей характеристикой системы является вероятность потери заявки. Отказ заявка получает, когда, как это было указано ранее, система находится в состоянии $\mathbf{S_{m+1}}$.Согласно (19.3) и (19.4)

$$\mathbf{P_{m+1}} = \rho^{\mathbf{m+1}}\mathbf{P_0} = \frac{\rho^{\mathbf{m+1}}(\mathbf{1}-\rho)}{\mathbf{1}-\rho^{\mathbf{m+2}}} \tag{19.5}$$

Относительная пропускная способность или вероятность быть обслуженным равна

$$\mathbf{q} = \mathbf{1} - \mathbf{P_{отк}} = \mathbf{1} - \frac{\rho^{\mathbf{m+1}}(\mathbf{1}-\rho)}{\mathbf{1}-\rho^{\mathbf{m+2}}} = \frac{\mathbf{1}-\rho^{\mathbf{m+1}}}{\mathbf{1}-\rho^{\mathbf{m+2}}} \tag{19.6}$$

Найдем среднее число заявок, находящихся в очереди $\bar{\mathbf{r}}$. Здесь не учитывается заявки, которые в данный момент обслуживаются.

С вероятностью $\mathbf{P_2}$ система находится в состоянии $\mathbf{S_2}$, значит в очереди стоит одна заявка.

С вероятностью $\mathbf{P_3}$ в очереди стоят две заявки.

С вероятностью $\mathbf{P_{m+1}}$ в очереди стоят $\mathbf{m}$ заявок.

Следовательно,

$$\bar{\mathbf{r}} = \mathbf{P_2 * 1 + P_3 * 2 + ... + P_{m+1} * m}$$

Согласно (19.4) выражаем все вероятности через $\mathbf{P_0}$, а затем подставляем его значение

$$\bar{\mathbf{r}} = \mathbf{1}\rho^{\mathbf{2}}\mathbf{P_0} + \mathbf{2}\rho^{\mathbf{3}}\mathbf{P_0} + ... + \mathbf{m}\rho^{\mathbf{m+1}}\mathbf{P_0} = \rho^{\mathbf{2}}\mathbf{P_0}[\mathbf{1} + \mathbf{2}\rho + \mathbf{3}\rho^{\mathbf{3}} + ... + \mathbf{m}\rho^{\mathbf{m-1}}] \tag{19.6a}$$

Выражение, стоящее в квадратных скобках, является производной от суммы $\sum\limits_{\mathbf{n=1}}^{\mathbf{m}}\rho^{\mathbf{n}}$, которая является геометрической прогрессией. Сумма прогрессии равна

$$\sum_{\mathbf{n=1}}^{\mathbf{m}}\rho^{\mathbf{n}} = \frac{\rho(\rho^{\mathbf{m}}-\mathbf{1})}{\rho-\mathbf{1}}$$

Следовательно,

$$\bar{\mathbf{r}} = \rho^2{}_0 \frac{\partial}{\partial\rho}\frac{\rho - \rho^{\mathbf{m}+1}}{1-\rho} = \rho^2 \mathbf{P}_0 \frac{(1-\rho)[1-(\mathbf{m}+1)\rho^{\mathbf{m}}] + \rho = \rho^{\mathbf{m}+1}}{(1-\rho)^2} =$$
$$= \rho^2 \frac{1-(\mathbf{m}+1)\rho^{\mathbf{m}} + \mathbf{m}\rho^{\mathbf{m}+1}}{(1-\rho^{\mathbf{m}+2})(1-\rho)} \tag{19.7}$$

Найдем теперь среднее число заявок, которые в текущий момент обслуживаются - $\bar{\varpi}$. Так как канал обслуживания один, то обслуживаться могут либо ноль заявок с вероятностью $\mathbf{P}_0$, либо одна заявка с вероятностью $\mathbf{1-P_0}$.

$$\bar{\varpi} = 0 * \mathbf{P}_0 + 1 * (1-\mathbf{P}_0) = \frac{\rho - \rho^{\mathbf{m}+2}}{1-\rho^{\mathbf{m}+2}} \tag{19.8}$$

Значит среднее число заявок, находящихся в системе, равно

$$\bar{\mathbf{u}} = \bar{\mathbf{r}} + \bar{\varpi} = \frac{\rho - (\mathbf{m}+2)\rho^{\mathbf{m}+2} + (\mathbf{m}+1)\rho^{\mathbf{m}+3}}{(1-\rho^2)(1-\rho)} \tag{19.9}$$

Найдем среднее время ожидания в очереди $\bar{\mathbf{t}}_{\mathbf{ож}}$. Пусть заявка пришла в некоторый момент времени. С вероятностью $\mathbf{P}_0$ канал в это время будет свободен и ей не придется стоять в очереди. С вероятностью $\mathbf{P}_1$ канал будет занят, очереди не будет и заявка станет первой в очереди. Ждать ей придется в течение среднего времени обслуживания предыдущей заявки, то есть в течение $\mathbf{1}/\mu$. Естественный вопрос о том, что мы не знаем сколько времени предыдущая заявка уже провела на обслуживании, неуместен. Экспоненциальный закон обслуживания делает время окончания обслуживания независимым от предыстории. Сколько времени заявка уже провела на обслуживании, не влияет на то, когда обслуживание закончится. С вероятностью $\mathbf{P}_2$ заявка будет второй в очереди. Ждать в очереди ей придется в среднем $\mathbf{2}/\mu$ и т. д. С вероятностью $\mathbf{P}_{\mathbf{m}+1}$ пришедшая заявка застает систему в состоянии $\mathbf{S}_{\mathbf{m}+1}$, когда очередь имеет предельную длину $\mathbf{m}$ и покидает систему необслуженной. Тогда

$$\bar{\mathbf{t}}_{\text{ож}} = \mathbf{P_1}\frac{\mathbf{1}}{\mu} + \mathbf{P_2}\frac{\mathbf{2}}{\mu} + ... + \mathbf{P_m}\frac{\mathbf{m}}{\mu}$$

Подставляя значения вероятностей, выраженных через $\mathbf{P_0}$ (см.(19.4)), получим

$$\bar{\mathbf{t}}_{\text{ож}} = \frac{\rho\mathbf{P_0}}{\mu}(\mathbf{1} + \mathbf{2}\rho + \mathbf{3}\rho^{\mathbf{2}} + ... + \mathbf{m}\rho^{\mathbf{m-1}})$$

Аналогичное, с точностью до множителя $\frac{\mathbf{1}}{\rho\mu}$, выражение было получено для величины $\bar{\mathbf{r}}$ (см. (19.6а)). Поэтому можно записать, выражение для среднего времени ожидания в виде

$$\bar{\mathbf{t}}_{\text{ож}} = \frac{\bar{\mathbf{r}}}{\rho\mu} = \frac{\bar{\mathbf{r}}}{\lambda} \tag{19.10}$$

Среднее время обслуживания определяется следующим образом. Заявка обслуживается с вероятностью, равной относительной пропускной способности $\mathbf{q}$. Среднее время обслуживания, при условии, что заявка обслуживается, равно $\mathbf{1}/\mu$. Поэтому среднее время обслуживании заявки (не каждая заявка обслуживается каналом, часть теряется) равно

$$\bar{\mathbf{t}}_{\text{об}} = \frac{\mathbf{q}}{\mu} \tag{19.11}$$

Итоговое среднее время пребывания заявки в системе равно

$$\bar{\mathbf{t}} = \bar{\mathbf{t}}_{\text{ож}} + \bar{\mathbf{t}}_{\text{об}} = \frac{\bar{\mathbf{r}}}{\lambda} + \frac{\mathbf{q}}{\mu} \tag{19.12}$$

Часто при расчете информационных систем необходимо определить объем буферного запоминающего устройства (БЗУ), предназначенного для сбора телеметрических отсчетов, поступающих по различным каналам связи с целью их обработки. БЗУ содержит $\mathbf{m}$ ячеек памяти и одну оперативную ячейку для подготовки отсчета для обработки. Предположим, что суммарный входной поток пуассоновский с интенсивностью λ. Выходной поток также пуассоновский с интенсивностью μ. Применим полученные ранее результаты к расчету БЗУ.

Основной характеристикой БЗУ является вероятность потерь, равная

$$\mathbf{P_{отк}} = \frac{\rho^{\mathbf{m+1}}(\mathbf{1}-\rho)}{\mathbf{1}-\rho^{\mathbf{m+2}}} \qquad (19.13)$$

Согласно теореме Шеннона о согласовании источника информации с каналом связи без помех следует выбрать $\lambda = \mu$. Рассмотрим формулу (19.13) при $\rho = \mathbf{1}$. Раскрываем неопределенность по правилу Лопиталя

$$\mathbf{P_{отк}} = \mathbf{lim}_{\rho \to 1} \frac{(\mathbf{m+1})\rho^{\mathbf{m}} - (\mathbf{m+2})\rho^{\mathbf{m+1}}}{-(\mathbf{m+2})\rho^{\mathbf{m+1}}} = \frac{\mathbf{1}}{\mathbf{m+2}}$$

При очень малых значениях допустимой вероятности потерь получаем обратно пропорциональное большое число ячеек памяти. Покажем, что при увеличении средней скорости обработки (уменьшении ρ) количество потребных ячеек памяти уменьшается. При $\rho = \mathbf{1/2}$

$$\mathbf{P_{отк}} = \frac{(\mathbf{1/2})^{\mathbf{m+1}} * \mathbf{1/2}}{\mathbf{1}-(\mathbf{1/2})^{\mathbf{m+2}}} = \frac{\mathbf{1}}{\mathbf{2^{m+2}}-\mathbf{1}}$$

Рассмотрим теперь характеристики одноканального СМО с ожиданием при неограниченной длине очереди $(\mathbf{m} \to \infty)$. Можно показать, что в СМО будут существовать предельные режимы только при $\rho < \mathbf{1}$, то есть тогда, когда интенсивность обслуживания больше интенсивности входного потока. При $\rho > \mathbf{1}$ длина очереди неограниченно растет.

Поэтому, при $\rho < \mathbf{1}$, найдем характеристики системы для бесконечной допустимой длины очереди. Вероятность отсутствия заявки получим из равенства (19.3)

$$\mathbf{P_0} = \mathbf{1} - \rho \qquad (19.14)$$

Вероятность потери заявки равна нулю, что и очевидно, так как любой заявке найдется место в очереди, и следует из (19.5). Вероятность быть обслуженным – относительная пропускная способность равна $\mathbf{q} = \mathbf{1}$

Средняя длина очереди согласно (19.7) равна

$$\bar{\mathbf{r}} = \frac{\rho^2}{1-\rho} \tag{19.15}$$

Среднее число заявок в системе согласно (19.9) равно

$$\bar{\mathbf{u}} = \frac{\rho^2}{1-\rho} + \rho = \frac{\rho}{1-\rho} \tag{19.16}$$

Среднее время пребывания в системе получаем согласно (19.12) и (19.15)

$$\bar{\mathbf{t}} = \frac{\rho^2}{\lambda(1-\rho)} + \frac{1}{\mu} = \frac{\rho}{\mu(1-\rho)} + \frac{1}{\mu} = \frac{1}{\mu(1-\rho)} \tag{19.17}$$

Обратите внимание, что средняя длина очереди и среднее время пребывания в системе катастрофически растут при приближении ρ к единице. Поэтому в бытовом плане нельзя строить СМО, рассчитывая на примерное равенство мощности обслуживания и мощности входного потока. Система развалится.

Этот эффект еще более усиливается в присутствии человеческого фактора. Наблюдения телефонных станций показывают, что с появлением признаков нехватки мощностей (приближение ρ к единице) рост очереди имеет вид функции от ρ в девятой степени.

Контрольные вопросы.

1. Нарисуйте граф одноканальной СМО с ожиданием. Обоснуйте принятые характеристики графа.

2. Приведите вывод выражений для установившихся значений вероятности простоя канала и вероятности потери заявки.

3. Приведите основания, принятые при выводе выражений для средней длине очереди в системе и для среднего времени пребывания заявок в системе.

4. Чему равна относительная пропускная способность системы при неограниченной длине очереди?

5. Как меняется средняя длина очереди в системе при увеличении отношения $\rho = \dfrac{\lambda}{\mu}$?

20. Буферное запоминающее устройство (БЗУ) с пуассоновским входным потоком и регулярным выходным потоком.

Рассмотрим одноканальную СМО с ограниченной длиной очереди, выполняющую роль буферного запоминающего устройства (БЗУ). БЗУ предназначено для хранения тех заявок из пуассоновского входного потока, которые при своем поступлении застали канал занятым. Их необходимо хранить пока не подойдет их очередь на обработку. Канал затрачивает на обработку каждой заявки одинаковое время, и обработанные заявки образуют регулярный выходной поток. Исключение составляют такие интервалы времени, когда в канале нет заявок для обработки, канал простаивает.

БЗУ можно представить в виде емкости , куда наливается входной поток со случайным расходом, а может выливаться через кран с фиксированным положением вентиля. Уровень жидкости в емкости поэтому меняется и основная задача заключается в том, чтобы уменьшить вероятность того, что жидкость перельется через край.

Рассматриваем состояние БЗУ через фиксированные промежутки времени τ , которые затрачиваютсь в БЗУ для обработки одной заявки. Сами моменты времени выбираются мгновением позже выхода очередной заявки и равны $\mathbf{t = n}\tau + \delta$, где δ - малая величина, $\mathbf{n}$ - целое.

Укажем на положения, принятые для составления системы уравнений. Вероятность прихода $\mathbf{k}$ заявок пуассоновского потока за время τ равно

$$\mathbf{P(k)} = \frac{(\lambda\tau)^{\mathbf{k}}}{\mathbf{k!}}\mathbf{e}^{-\lambda\tau} \qquad (20.1)$$

$\mathbf{m}$ - число ячеек БЗУ.

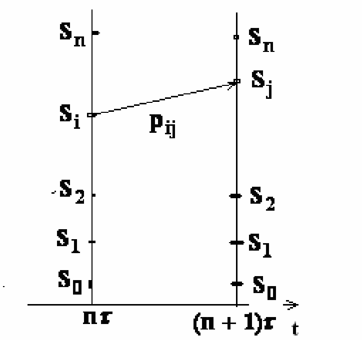
Состояния БЗУ обозначим $\mathbf{S_i}$, где $\mathbf{i}$ число заявок в устройстве.

Рассматриваем установившиеся вероятности состояний БЗУ, поэтому

$$\mathbf{Q_i(n}\tau\mathbf{)} = \mathbf{Q_i[(n+1)}\tau\mathbf{]} \qquad (20.2)$$

Состояние $\mathbf{S_0}$ - в системе нет заявок. Канал обработки простаивает.

Состояние $\mathbf{S_n}$ - все ячейки БЗУ заполнены. Заявки, заставшие систему в этом состоянии, будут утеряны.



Условные вероятности перехода
Рис. 20.1

На Рис.20.1 показаны состояния БЗУ в соседние моменты времени. $\mathbf{p_{ij}}$ - условная вероятность перехода системы из состояния $\mathbf{S_i}$ (в системе $\mathbf{i}$ заявок) в состояние $\mathbf{S_j}$ (в системе $\mathbf{j}$ заявок). Полагаем, что все условные вероятности от времени не зависят.

В системе в момент времени $\mathbf{n}\tau$ может находиться произвольное число заявок. За интервал τ может появиться случайное число заявок, вероятность которого определяется (20.1). Уйти из системы за этот интервал может только одна заявка, если хотя бы одна заявка в системе есть.

Составим выражения для условных вероятностей перехода .

Условная вероятность перехода $\mathbf{p_{00}}$- вероятность перехода из состояния $\mathbf{S_0}$ в начале рассматриваемого интервала времени в то же состояние $\mathbf{S_0}$ в конце временного интервала. Это означает, что в начале в системе не было заявок, поэтому из системы ничего не было передано. За временной интервал в систему ничего не пришло, и система осталась без заявок. Условная вероятность этого случая (условие – в системе не было заявок) равна согласно (20.1)

$$\mathbf{P(0) = e^{-\rho}} \tag{20.3}$$

Здесь обозначено

$$\rho = \lambda\Delta \tag{20.4}$$

которое имеет тот же смысл отношения среднего количества пришедших заявок к количеству, равному единице, выходных заявок.

Полная вероятность рассматриваемого случая перехода равна

$$\mathbf{P_{00} = Q_0 p_{00}} \qquad (20.5)$$

Условная вероятность перехода $\mathbf{p_{01}}$ (переход из состояния $\mathbf{S_0}$ в $\mathbf{S_1}$, что означает, что в системе вначале не было заявок, поэтому ничего передано не было и за рассматриваемый интервал пришла одна заявка), согласно (20.1)

$$\mathbf{p_{01} = P(1) = \rho e^{-\rho}} \qquad (20.6)$$

Полная вероятность этого случая равна

$$\mathbf{P_{01} = Q_0 p_{01}} \qquad (20.7)$$

Условная вероятность $\mathbf{p_{02}}$ отличается тем, ч то в систему придут две заявки, поэтому

$$\mathbf{p_{02} = P(2) = \frac{\rho^2}{2!} e^{-\rho}} \qquad (20.7)$$

Полная вероятность равна

$$\mathbf{P_{02} = Q_0 p_{02}} \qquad (20.8)$$

При вычислении условной вероятности $\mathbf{p_{10}}$ следует учесть, что в системе была одна заявка, она была передана, и в систему ничего не пришло. Поэтому условная вероятность равна

$$\mathbf{p_{10} = P(0) = e^{-\rho}} \qquad (20.9)$$

Полная вероятность равна

$$\mathbf{P_{10} = Q_1 p_{10}} \qquad (20.10)$$

Условная вероятность $\mathbf{p_{11}}$ отличается тем, что была передана одна заявка и пришла одна заявка

$$\mathbf{p_{11} = P(1) = \rho e^{-\rho}} \qquad (20.11)$$

Полная вероятность равна

$$P_{11} = Q_1 p_{11} \qquad (20.12)$$

Остальные случаи перехода из состояния S_1 очевидны, кроме перехода в состояние S_n, которое будет рассмотрено позднее.

Рассмотрим переход из состояния S_2 в состояние S_0, что означает - из системы было передано две заявки. Это невозможно, больше одной заявки за интервал из системы передать нельзя. Это справедливо для всех состояний

$$P_{ij} = 0 \text{ если } j < i - 1 \qquad (20.13)$$

Переход из состояния S_2 в состояние S_1 означает, что из системы была передана одна заявка, и ничего не пришло.

Имеем $P_{21} = Q_2 p_{21} = Q_2 e^{-\rho}$. (20.14)

Остальные случаи очевидны. Рассмотрим случаи перехода в состояние S_n. В этом состоянии все ячейки БЗУ заполнены и очередные заявки, сколько бы их не пришло, получат отказ и система останется в состоянии S_0. Поэтому

$$p_{i,n} = \sum_{j=n-i+1}^{\infty} P(j) \text{ при } i \neq 0 \qquad (20.15)$$

$$p_{0,n} = \sum_{j=n}^{\infty} P(j) \qquad (20.15а)$$

Составим теперь выражения для установившихся вероятностей состояний. Учитывая вероятные пути, которыми можно прийти в каждое состояние в конце рассматриваемого интервала времени, учитывая ограничение (20.13), учитывая состояние стационарности (20.2), получим

$$\begin{aligned}
&P_0 = P_0 p_{00} + P_1 p_{10} \\
&P_1 = P_0 p_{01} + P_1 p_{11} + P_2 p_{21} \\
&P_2 = P_0 p_{02} + P_1 p_{12} + P_2 p_{22} + P_3 p_{32} \\
&\ldots\ldots\ldots\ldots \\
&P_n = 1 - (P_0 + P_1 + P_2 + \ldots + P_{n-1})
\end{aligned} \qquad (20.16)$$

Обратите внимание, что последнее уравнение заменяет собой сложные соотношения (20.15), (20.15а). Это всегда рекомендуется использовать. Далее в систему подставляются конкретные значения условных вероятностей перехода $\mathbf{p_{ij}}$, и находится решение неоднородного линейного алгебраического уравнения.

Например, при $\mathbf{n = 2}$ система уравнений имеет вид

$$\begin{aligned}
&\mathbf{P_0 = P_0 e^{-\rho} + P_1 e^{-\rho}}\\
&\mathbf{P_1 = P_0 \rho e^{-\rho} + P_1 \rho e^{-\rho} + P_2 e^{-\rho}}\\
&\mathbf{P_2 = 1 - (P_1 + P_2)}
\end{aligned}$$

Определим вероятность потери заявок. Вероятность простоя системы или вероятность того, что из системы не будет передано ни одной заявки, равна $\mathbf{P_0}$. Тогда вероятность того, что система за один интервал передаст одну заявку, равна $\mathbf{1 - P_0}$. За время $\mathbf{T}$ в систему придут в среднем $\lambda\mathbf{T}$ заявок. Следовательно, в среднем будет потеряно $\lambda\mathbf{T} - \mathbf{(1 - P_0)}\frac{\mathbf{T}}{\tau}$ заявок. Вероятность потерь оценим как долю потерянных заявок к общему числу заявок

$$\mathbf{P_{пот}} = \frac{\lambda\mathbf{T} - \mathbf{(1 - P_0)}\frac{\mathbf{T}}{\tau}}{\lambda\mathbf{T}} = \frac{\lambda - \mathbf{(1 - P_0)}\frac{\mathbf{1}}{\tau}}{\lambda} = \frac{\rho - \mathbf{(1 - P_0)}}{\rho} \tag{20.17}$$

На Рис 20.2 показана зависимость вероятности потерь от количества ячеек в БЗУ и относительной мощности системы ρ. Для получения результата каждому значению $\mathbf{n}$ соответствовала своя система уравнений (20.16), которая решалась для каждого значения ρ.

Контрольные вопросы

1. Запишите вид системы уравнений для определения вероятностей состояний системы.

2. Опишите подход к определению вероятности потери заявки в БЗУ.

21. Поиск неисправностей типа «обрыв» и «короткое замыкание»

Рассмотрим системы с некоторыми неисправностями. Из всех возможных неисправностей элементов системы можно в рамках использования аппарата логической алгебры выделить две большие группы:

1. Неисправность типа «обрыв», когда выход элемента при любой комбинации входов равен нулю.

2. Неисправность типа «короткое замыкание», когда выход элемента при любой комбинации входов равен единице.

Сначала рассмотрим неисправность первого типа. Очевидно, что логический элемент в этом случае может быть описан в виде

$$\mathbf{Y=NY^*} \tag{21.1}$$

Где **Y*** - выход, вычисленный для исправного элемента ,

N – показатель неисправности,

N=1 - для исправного элемента,

N=0 – для неисправности типа «обрыв».

При неисправности типа «короткое замыкание элемент может быть описан в виде

$$\mathbf{Y=Y^*+K} \tag{21.2}$$

Здесь **К** – показатель неисправности «короткое замыкание»,

К=0 – для исправного элемента,

К=1 – для неисправности типа «короткое замыкание».

Если возможны оба вида неисправности одновременно

$$\mathbf{Y=NY^*+K}$$

Для поиска неисправности типа «короткое замыкание» автор в настоящее время может предложить следующий способ. Для каждого элемента подбираются такие входы, при которых у исправного элемента выход должен быть равен нулю. Если реальный выход равен единице, следовательно имеет место неисправность «короткое замыкание».

Перейдем к неисправности типа «обрыв». Рассмотрим схему, показанную на рис 21.1.

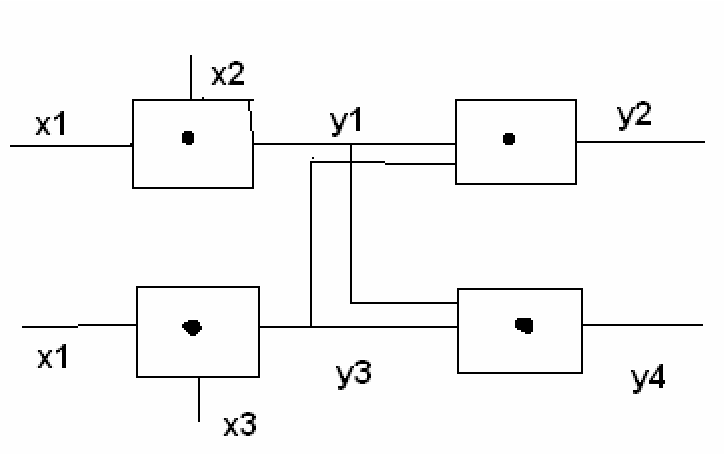


Рис.21.1

Для этой схемы система уравнений с учетом возможных неисправностей типа «обрыв» имеет вид

$$\begin{aligned} &y_1=N_1x_1x_2\\ &y_2=N_2y_1y_3\\ &y_3=N_3x_1x_3\\ &y_4=N_4y_1y_3 \end{aligned} \qquad (21.3)$$

Исключая внутренние выходы из правых частей, получим

$$\begin{aligned} &y_1 = N_1x_1x_2\\ &y_2 = N_1N_2N_3x_1x_2x_3\\ &y_3 = N_3x_1x_3\\ &y_4 = N_1N_3N_4x_1x_2x_3 \end{aligned} \qquad (21.4)$$

Покажем возможности использования полученного аппарата в задачах обнаружения неисправностей – технической диагностики. Процесс обнаружения неисправностей состоит в том, что подается некоторый набор входных сигналов и по результатам измеренных выходных сигналов судят о состоянии системы.

Необходимо определить, какие входы необходимо подавать и какие выходы измерять, чтобы обнаружить все возможные неисправности.

Для примера рассматриваем простейший случай, когда система состоит из блоков, реализующих логическое умножение. В этом случае, как это видно из (21.4) необходимо подавать все входы, иначе ряд выходов будут равны нулю независимо от состояния элемента.

Определим теперь, какие выходы необходимо измерять для определения неисправностей. Составляется таблица неисправностей (таблица 21.1), где указываются значения выходов при различных неисправностях системы.

Таблица 21.1

	n_0	n_1	n_2	n_3	n_4	n_{12}	n_{13}	…….	n_{1234}
y_1	1	0	1	1	1	0	0		0
y_2	1	0	0	0	1	0	0		0
y_3	1	1	1	0	1	1	0		0
y_4	1	0	1	0	0	0	0		0

Значения в таблице вычислены в соответствии с уравнениями (21.4).

n_0 – исправное состояние системы (N_1=N_2=N_3=N_4=1),

n_i – неисправность i – го элемента (N_i=0),

n_{ij} – кратная неисправность (N_i=N_j=0).

Таблица неисправности может быть составлена либо путем анализа математического описания системы, либо путем моделирования неисправностей, либо путем анализа структурной схемы и т. п.

Результат измерения выходов есть кодовая комбинация, по которой можно различить состояние системы. Если кодовые комбинации для ряда состояний совпадают, состояния называются неразличимыми. Например, состояния n_1 и n_{12} в таблице 21.1.

Обычно таблица неисправностей содержит все выходные сигналы, которые могут быть измерены, и число этих сигналов может оказаться избыточным. Найдем минимальное число выходов, по которым можно составить кодовые комбинации для всех различимых состояний. Решение этой задачи дадим на примере таблицы 21.1.

Чтобы отличить состояние n_0 от состояния n_1, достаточно измерить один выход, который имеет разный исход для этих состояний. Разные значения в столбцах n_0 и n_1 имеют выходы y_1, y_2,y_4. Следовательно, диагностический тест t_{01}, отличающий состояние n_0 от состояния n_1, заключается в измерении выхода y_1 ИЛИ выхода y_2 ИЛИ выхода y_4, что можно записать в виде

$$\mathbf{t_{01} = y_1 + y_2 + y_4} \qquad (21.5)$$

Аналогично, для отличия состояний n_0 от n_2 необходимо измерить

$$\mathbf{t_{02} = y_2} \tag{21.6}$$

а для отличия n_1 от n_2 необходимо измерить

$$\mathbf{t_{12} = y_2 + y_4} \tag{21.7}$$

и т. д.

Таблица 21.2

	y_1	y_2	y_3	y_4
t_{01}	1	1	0	1
t_{02}	0	1	0	0
t_{03}	0	1	1	1
t_{04}	0	0	0	1
t_{12}	1	0	0	1
t_{13}	1	0	1	0
t_{14}	1	1	0	0
t_{23}	0	0	1	1
t_{24}	0	1	0	1
t_{34}	0	1	1	0

Составленные таким образом выражения t_{ij} записываются в виде таблицы различимости. Таблица различимости для рассматриваемого примера в случае одиночных неисправностей приведена в таблице 21.3. Наличие слагаемого отмечается единицей, отсутствие – нулем.

Для решения задачи диагностики – определения места неисправности необходимо попарно различить все состояния, то есть реализовать все тесты. Следовательно, полный диагностический тест **T** есть произведение всех тестов $\mathbf{t_{ij}}$ из таблицы различимости.

Необходимо измерить t_{01} И t_{02} И t_{03} И т.д. Следовательно

$$\begin{aligned}\mathbf{T} &= \mathbf{(y_1 + y_2 + y_4)y_2(y_2 + y_3 + y_4)y_4(y_1 + y_4)(y_1 + y_3)(y_1 + y_2)} \cap \\ &\cap \mathbf{(y_3 + y_4)(y_2 + y_4)(y_2 + y_3)}\end{aligned} \tag{21.8}$$

Преобразуем полученное выражение из КНФ в ДНФ, производя одновременно минимизацию выражения с помощью операции сокращения

$$\mathbf{T = y_2 y_4 (y_1 + y_3) = y_1 y_2 y_4 + y_2 y_3 y_4} \tag{21.9}$$

Этот результат в соответствии со смыслом операций логического умножения и сложения означает, что для определения состояния системы необходимо измерить выходы y_1 И y_2 И y_4 ИЛИ выходы y_2 И y_3 И y_4. Получены два диагностических теста, оба они в нашей постановке равноправны. Действительно, по таблице (21.2) легко убедиться, что результаты измерения трех выходов y_1 , y_2 , y_4 дают различные кодовые комбинации для всех одиночных неисправностей. Аналогичными свойствами обладает второй тест y_2 , y_3 , y_4.

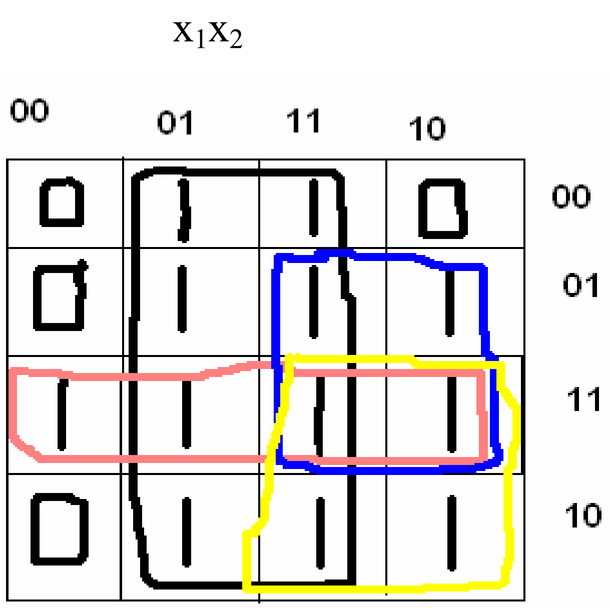
Таким образом, диагностический тест составляется по таблице различимости, каждая строка которого есть множитель теста, составленного в виде КНФ. Множитель представляет собой сумму тех переменных, которые равны в этой строке единице. После преобразования КНФ в ДНФ и минимизации полученного выражения каждое слагаемое ДНФ представляет набор выходов, достаточный для диагностики системы. Все наборы равноправны, выбирается какой – то один. Задача выбора оптимального теста, выбора оптимальной последовательности проверки и т.п. здесь не рассматривается.

Задача контроля отличается от задачи диагностики тем, что определяется, исправна ли система. Место неисправности не определяется. Следовательно, контрольный тест должен отличить исправное состояние n_0 от всех других. В случае одиночных неисправностей согласно таблице 21.2 контрольным тестом будет

$$T_k = (y_1 + y_2 + y_4)y_2(y_2 + y_3 + y_4)y_4 = y_2y_4 \qquad (21.10)$$

Действительно, из таблицы 21.1 видно, что измерение двух выходов y_2 и y_4 дает возможность отличить исправное состояние системы от неисправного.

Далее рассмотрим вопрос перевода КНФ в ДНФ, который в случае большой размерности системы не так прост.



В начале рассмотрим возможность использования матриц Карно. В качестве примера рассмотрим диагностический тест в виде КНФ

$$\mathbf{T = (x_1 + x_2 + x_4)(x_2 + x_3 + x_4)(x_1 + x_2 + x_3)} \qquad (\quad .11)$$

Условимся, что клетки матрицы Карно, занятые множителем КНФ, определяются следующим образом:

- переменная, входящая в множитель в прямом виде имеет соответствующую координату клетки, равную нулю,

- клетки помечаются нулем.

Известно, что такое же выражение, но в виде ДНФ, занимает на матрице Карно все свободные от КНФ клетки. Эти клетки будем помечать единицами, а координаты переменной, входящих в ДНФ в прямом виде считать равными единице.

Три множителя выражения (21.11) заняли на матрице Карно четыре клетки и помечены нулями. Остальные клетки помечены единицами и являются слагаемыми ДНФ. Клетки, обведенные черной границей, дают слагаемое x_2. Клетки, обведенные красной границей, дают слагаемое x_3x_4. Клетки, обведенные синей и желтой границами, дают слагаемые x_1x_4 и x_3x_4. Таким образом тест в виде ДНФ равен

$$\mathbf{T = x_2 + x_3x_4 + x_1x_4 + x_3x_4} \qquad (21.12)$$

Заметим, что слагаемые ДНФ не содержат инверсий выходов, что противоречило бы физическому смыслу задачи.

Преобразование (.11) в ДНФ простым перемножением подтверждает результат (.12).

Перевод КНФ в ДНФ при большом числе выходных сигналов – задача очень трудоемкая. Кроме того при переводе необходимо реализовать минимизацию выражения. Например из двух слагаемых теста $\mathbf{y_1y_2 + y_1y_2y_3}$ следует оставить только первое, так как проверка выхода $\mathbf{y_3}$ очевидно является лишней. Устранение лишних проверок произойдет в процессе минимизации ДНФ. Поскольку исходные и конечные логические выражения не должны содержать инверсий, единственной операцией минимизации является операция поглощения. Из большого многообразия переводов КНФ в ДНФ укажем следующий способ.

Любая логическая функция может быть представлена с помощью формулы Шеннона в виде

$$\mathbf{F(y_1, y_2, ... y_n) = y_1 F(1, y_2, ... y_n) + \overline{y_1}\ F(0, y_2, ... y_n)} \qquad (21.13)$$

Под параметром $\mathbf{y_1}$ можно принять любой выход. Из (21.13), например, следует:

$$\mathbf{(y_1 + y_2)(y_2 + y_3) = y_1(1 + y_2)(y_2 + y_3) + \overline{y}_1(0 + y_2)(y_2 + y_3)} \qquad (21.14)$$

Формулу (21.13) можно использовать для перевода КНФ в ДНФ. После применения (21.13) к $\mathbf{F(y_1, ... y_n)}$, заданного в виде КНФ, получаем два слагаемых. Далее (21.13) применяется к каждому слагаемому и т. д. Однако (21.13) содержит инверсию переменных, в то время как исходная функция по смыслу задачи инверсий содержать не может. Поэтому преобразуем (21.13) к более удобному виду.

Так как 1+y=1, то $\mathbf{F(1, y_2, ... y_n)}$ есть произведение только тех множителей $\mathbf{F(y_1, y_2, ... y_n)}$, которые не содержат переменную $\mathbf{y_1}$.

Выражение $\mathbf{F(0, y_2, ... y_n)}$ можно представить как произведение всех множителей исходной КНФ, куда не входило $\mathbf{y_1}$, то есть $\mathbf{F(1, y_2, ... y_n)}$, умноженное на произведение тех множителей, куда входили $\mathbf{y_1}$, и где теперь сделана замена $\mathbf{y_1 = 0}$. Обозначим произведение последних множителей $\mathbf{Q(y_2, ... y_n)}$, получим

$$\mathbf{F(0, y_2, ... y_n) = F(1, y_2, ... y_n) Q(y_2, ... y_n)} \qquad (21.15)$$

Например в (21.14)

$$\mathbf{F(1, y_2, y_3) = (y_2 + y_3)}$$

$$\mathbf{F(0, y_2, y_3) = y_2(y_2 + y_3)}$$

$$\mathbf{Q(y_2, y_3) = y_2}$$

Подставим (21.15) в (21.13)

$$\mathbf{F(y_1, y_2, ... y_n) = y_1 F(1, y_2, ... y_n) + \overline{y}_1 Q(y_2, ... y_n) F(1, y_2, ... y_n)} \qquad (21.16)$$

Так как $\mathbf{Q}$ есть логическая функция, принимающая два значения 0 и 1, то

$$\mathbf{y_1 + \overline{y}_1 Q = y_1 + Q}, \qquad (21.17)$$

Что может быть подтверждено вычислениями в табл.21.3

Таблица 21.3

$\mathbf{y}$	$\mathbf{Q}$	$\mathbf{\overline{y}}$	$\mathbf{y + \overline{y}Q}$	$\mathbf{y + Q}$
0	0	1	0	0
0	1	1	1	1
1	0	0	1	1
1	1	0	1	1

Учитывая в (21.16) результаты (21.15) и (21.17), получим

$$\mathbf{F(y_1, y_2, ...y_n) = y_1F(1, y_2, ...y_n) + F(0, y_2, ...y_n)} \qquad (21.18)$$

Инверсия переменной исчезла, что соответствует смыслу задачи. Применяя последовательно эту формулу для следующей переменной к каждому слагаемому (21.18) получим выражение в виде ДНФ.

Применим к КНФ в виде таблицы различимости (табл. 21.4) способ (21.18). Получим

$$\mathbf{F(y_1, ...y_n) = y_1F(1, ...y_n)\{tab21.5\} + F(0, ...y_n)\{tab21.6\}} \qquad (21.19)$$

Таблица 21.4

y_1	y_2	y_3	y_4	y_5	y_6
0	1	0	0	1	0
0	1	1	1	0	0
1	1	0	0	0	1
1	0	1	0	0	0
0	0	1	0	0	1
1	0	0	1	0	0
1	0	0	0	1	0

Таблица 21.5

y_2	y_3	y_4	y_5	y_6
1	0	0	1	0
1	1	1	0	0
0	1	0	0	1

Таблица 21.6

y_2	y_3	y_4	y_5	y_6
1	0	0	1	0
1	1	1	0	0
1	0	0	0	1
0	1	0	0	0
0	1	0	0	1
0	0	1	0	0
0	0	0	1	0

Теперь к каждой таблице 21.5 и 21.6 применяем процедуру (21.18). Аналогично поступаем с каждой полученной в дальнейшем таблицей пока не придем к выражению в форме ДНФ. При этом рекомендуется следующая приоритетность действий.

1. Если в таблице имеется столбец, содержащий одни нули, столбец просто вычеркивается, так как таблица соответствующий параметр не содержит
2. Если имеется строка, содержащая одну единицу в столбце $\mathbf{y_i}$, это означает, что параметр $\mathbf{y_i}$ входит в ДНФ простым множителем и поглощает все скобки, содержащие этот параметр. Следовательно, таблица равна произведению

$$\mathbf{y_i F(y_k, ... y_n)_{y_i = 1}}$$

Напоминаем, что второй множитель образуется из исходной таблицы вычеркиванием столбца y_i и удалением строк, где $y_i = 1$. Эту операцию поглощения пропускать нельзя, так как это единственная возможность минимизировать итоговую КНФ.

3. Если действия по пунктам 1 и 2 невозможны, для последующего разложения выбирается параметр, имеющий наибольшее число единиц в столбце.

Получим, что первое слагаемое в (21.19) равно

$$y_1y_2(y_3+y_4)+y_1\begin{vmatrix} y_3 & y_4 & y_5 & y_6 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{vmatrix} = y_1y_2y_3+y_1y_2y_4+y_1y_5\begin{vmatrix} y_3 & y_4 & y_6 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} =$$

$$= y_1y_2y_3+y_1y_2y_4+y_1y_5(y_3+y_4y_6) = y_1y_2y_3+y_1y_2y_4+y_1y_3y_5+y_1y_4y_5y_6$$

Второе слагаемое в (21.19) равно

$$\begin{vmatrix} y_2 & y_3 & y_4 & y_5 & y_6 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{vmatrix} = y_3y_4y_5(y_2+y_6) = y_2y_3y_4y_5+y_3y_4y_5y_6$$

Таким образом, окончательное выражение КНФ содержит шесть слагаемых без инверсий переменных.

Контрольные вопросы

1. Что называется неисправностями типа «обрыв» и «короткое замыкание»?

2.Как составляются таблица неисправностей и таблица различимости?

3. Как составляются диагностические и контрольные тесты?

4. Какие способы минимизации выражений для диагностических тестов Вы знаете?

22. Надежность программных средств [3]

Было показано на примере разработки программ противоракетной обороны США (СОИ), что при существующем уровне развития методов, технологий и средств автоматизации проектирования есть предельные масштабы и объемы комплексов программ, при которых невозможно обеспечить при допустимых затратах необходимую надежность их функционирования и безопасность применения. Очень высокие требования к качеству программ принципиально не могут быть выполнены, вследствие реальных ограничений всех видов ресурсов: бюджета, длительности разработки, характеристик вычислительных систем и квалификации специалистов. Применение таких систем при недостаточной их надежности и безопасности теряет смысл и может становиться не только бесполезным, но и опасным.

Для систематической, координированной борьбы с угрозами надежности должны проводиться исследование конкретных факторов, влияющих на качество функционирования и безопасность применения программ со стороны, реально существующих и потенциально возможных дефектов в создаваемых комплексах программ. В каждом проекте должен целенаправленно разрабатываться скоординированный комплекс методов и средств обеспечения заданной надежности функционирования ПС при реально достижимом уровне снижения дефектов и ошибок разработки.

Контроль надежности и безопасности создаваемых и модифицируемых программ должен сопровождать весь жизненный цикл ПС посредством специальной, достаточно эффективной технологической системы обеспечения их качества.

Для обнаружения и устранения ошибок проектирования все этапы разработки и сопровождения ПС должны быть поддержаны методами и средствами систематического, автоматизированного тестирования и испытаний.

22.1 Надежность функционирования программных средств

Надежность технических систем определяется в основном двумя факторами: надежностью компонент и дефектами в конструкции, допущенными при проектировании и изготовлении. Надежность сложных программных средств определяется этими же факторами, однако доминирующими являются дефекты и ошибки проектирования, так как физическое хранение программ характеризуется очень высокой надежностью. Программа любой сложности и назначения при строго фиксированных исходных данных и абсолютно надежной аппаратуре исполняется по однозначно определенному маршруту и дает на выходе строго определенный результат. Однако случайное изменение исходных данных и накопленной при обработке информации, а также множество условных переходов в программе создают огромное число различных маршрутов исполнения каждого сложного ПС. Источниками ненадежности являются непроверенные сочетания исходных данных, при которых функционирующее ПС дает неверные результаты и отказы.

При применении понятия к программным средствам следует учитывать особенности и отличия этих объектов от традиционных технических систем, для которых в основном разрабатывалась теория надежности:

- не для всех видов программ применимы понятия теории надежности,

- основными факторами, определяющими надежность программ, являются дефекты и ошибки проектирования и разработки, и второстепенное значение имеет физическое разрушение программных компонент при внешних воздействиях,

- относительно редкое разрушение программных компонент и необходимость их физической замены приводит к принципиальному изменению понятий сбоя и отказа программ,

- для повышения надежности программ особое значение имеет сокращение длительности восстановления и преобразования отказов в кратковременные сбои путем введения в программные средства временной, программной и информационной избыточности,

- методы форсированных испытаний надежности систем путем физического воздействия на их компоненты неприменимы для программных средств, и их следует заменять методами форсированного воздействия информационных потоков внешней среды.

Поэтому основные существующие понятия теории надежности нужно адаптировать к оценке качества программ и развивать эту теорию в особом направлении – надежность программных средств.

Рекомендуется шесть основных характеристик качества ПС, каждая из которых детализируется своими субхарактеристиками:

Функциональная пригодность определяется пригодностью для применения, точностью, защищенностью, способностью к взаимодействию и согласованностью со стандартами проектирования.

Надежность рекомендуется характеризовать степенью отсутствия ошибок, устойчивостью к ошибкам и перезапускаемостью.

Применимость предлагается описывать понятностью, обучаемостью и простотой использования.

Эффективность рекомендуется характеризовать ресурсной и временной экономичностью.

Сопровождаемость описывается качеством документации, удобством для анализа, изменяемостью, стабильностью и тестируемостью.

Переносимость предлагается отражать адаптируемостью, структурированностью, замещаемостью и внедряемостью.

Различие между ожидаемыми и полученными результатами функционирования программ могут быть следствием не только ошибок в созданных программах и данных, но и системных ошибок в требованиях спецификаций, явившихся исходной базой для создания ПС. Установить ошибочность исходных данных и спецификаций еще труднее, чем обнаружить ошибки в программах и данных, так как принципиально отсутствуют формализованные данные, которые можно использовать как эталонные, и их заменяют неформализованные представления заказчиков и разработчиков.

Степень влияния всех внутренних дестабилизирующих факторов, а также некоторых внешних угроз, на надежность ПС определяется в наибольшей степени качеством технологии проектирования, разработки, сопровождения и документирования ПС и их основных компонент. При ограниченных ресурсах на разработку, для достижения требований по надежности необходимо управление обеспечением качества в течение всего цикла создания программ и данных.

Среди методов обеспечения надежности ПС можно выделить следующие:

- создание программных модулей и компонентов гарантированного качества,
- предотвращение дефектов проектирования за счет эффективных технологий и средств автоматизации всего жизненного цикла программ и баз данных,
- обнаружение и устранение различных дефектов и ошибок путем систематического тестирования на всех этапах жизненного цикла,
- проверка достигнутого качества и надежности функционирования и сертификация системы перед передачей в регулярную эксплуатацию,
- обеспечение технической поддержки пользователя в процессе эксплуатации программного продукта.

<u>Предотвращение ошибок</u> и улучшение показателей создания ПС обеспечивается применением современных технологий, CASE систем автоматизированного проектирования и языков четвертого поколения (4GL), что способно снизить трудоемкость сложных ПС в несколько раз, заметно

повысить надежность проектирования и сократить длительность проектирования с 2-3 лет до нескольких месяцев.

Базовым принципом современных методов и технологий создания ПС является использование отработанных технических решений на различных аппаратных и операционных платформах. В настоящее время, по некоторым оценкам, только 10-15% прикладных программ создается вновь, в то время как основная часть программных средств переносится с других проектов и платформ гарантированного качества.

Языки 4GL разных видов имеют общую цель – ускорение разработок, повышение их надежности и снижение требований к уровню квалификации разработчика.

Для обнаружения и устранения ошибок проектирования все этапы разработки и сопровождения ПС должны быть поддержаны методами и средствами систематического автоматизированного тестирования. Тестирование – основной метод измерения качества, определения корректности и реальной надежности функционирования программ на любых этапах разработки.

Большое значение имеет учет особенностей тестирования сложных программ, которые отличают этот процесс от традиционного, применяемого для проверки аппаратуры и других технических систем. С этой позиции, основными особенностями тестирования являются:

- отсутствие достоверного эталона – программы, которому должны соответствовать результаты тестирования,

- высокая сложность комплексов ПС и принципиальная невозможность построения полных комплектов тестовых наборов,

- относительно невысокая степень формализации критериев качества тестирования и достигаемых при этом корректности и надежности.

Недооценка необходимости планомерного тестирования в процессе разработки проекта приводит к резкому возрастанию затрат на выявление и исправление ошибок в процессе эксплуатации.

Отсутствие гарантии достижения в процессе создания ПС абсолютной надежности их функционирования заставляет искать дополнительные методы и средства повышения надежности ПС. Для этого разрабатываются и применяются методы введения в них временной, информационной и программной избыточности.

При использовании зарубежных ПС следует иметь в виду, что, в принципе, в них возможны как случайные, так и намеренные искажения вычислительного процесса, программ и данных. Их применение в ответственных отечественных ПС требует дополнительного контроля качества и спецработ по обеспечению надежности. Иногда требуется и разрешение вышестоящих организаций. Вышеизложенное верно и при использовании программного продукта конкурирующих фирм и организаций.

22.2 Принципы и этапы тестирования

В данном подразделе описываются основные методы тестирования относительно простых программ. К ним относятся отдельные программные модули (ПМ) или их небольшие группы, решающие достаточно простую функциональную задачу. Невысокая размерность ПМ обеспечивает их обозримость и возможность детального анализа функций, структуры и процесса решения задач.

Методы и стратегия тестирования. При детерминированном тестировании задаются конкретные совокупности исходных данных, которым соответствуют заранее полученные эталонные результаты. В процессе тестирования на выходе и в промежуточных точках полученные результаты сопоставляются с эталонными, что позволяет не только обнаружить ошибку, но, в ряде случаев, и локализовать ее.

Стохастические методы тестирования предполагают использование исходных данных в виде совокупности случайных величин с известными распределениями. Им должны быть сопоставлены эталонные распределения результатов, которые в совокупности образуют тесты. Оценка результатов тестирования производится по степени соответствия полученных распределений и их параметров эталонным распределениям.

Тестирование потоков управления. Тестирование потоков управления (структуры программы) является начальным этапом, так как при некорректной структуре возможны наиболее грубые искажения выходных результатов. Тестирование состоит в проверке корректности последовательностей передач управления и формирования маршрутов исполнения программы при различных исходных данных.

Тестирование потоков данных можно разделить на два этапа. Первый этап тестирования состоит в анализе и обработке данных, определяющих значение предиката в операторах выработки логических решений. Эти решения влияют на маршруты обработки информации, что сближает этот метод тестирования с предыдущим. Второй этап тестирования обработки данных состоит в проверке вычислений по формулам численных значений результатов в зависимости от числовых значений исходных данных. В качестве эталонов используются результаты ручных или автоматизированных расчетов по тем же формулам. Возможно также проведение сравнений с аналитическими оценками результатов.

Искусство программиста состоит в создании модулей с относительно простой структурой. Относительная простота ПМ позволяет детально анализировать внутреннюю структуру и любой маршрут исполнения программы. Это обеспечивает возможность реализации двух стратегий отладки: от структуры или от данных. Этим двум стратегиям соответствуют два метода тестирования программ: метод анализа потоков управления и метод анализа потоков данных. Эти методы дополняют друг друга. Тестированию помогает хорошо отработанная блок-схема программы.

При первой стратегии (Рис. 22.1) за основу принимается структура ПМ, построенная по тексту программы в виде направленного графа. В графе программы по некоторым критериям выделяются и упорядочиваются маршруты исполнения программы и условия – предикаты, при которых они могут быть реализованы. Эти условия используются для подготовки тестовых наборов, каждый из которых должен реализоваться по маршруту, принятому за эталон при подготовке теста. Отклонение исполнения теста от первоначально выбранного маршрута рассматривается как ошибка, локализация которой может быть либо в первичной структуре ПМ, либо в реализации конкретного маршрута при данном тесте на входе.

Схема первой стратегии отладки
Рис.23.1

Схема второй стратегии отладки
Рис.23.2

Для каждого маршрута затем проверяется процесс обработки данных. Затем оценивается достаточность выполненного тестирования по степени покрытия исходного графа программы проверенными маршрутами. Завершается отладка при полном покрытии графа протестированными маршрутами. В случае обнаружения ветвей графа, выход на которые невозможен, следует проверить непротиворечивость условий выхода на этот участок и, возможно, удалить лишние ветви графа.

Данная стратегия имеет некоторое преимущество при тестировании логических программ с малой долей вычислений.

Вторая стратегия (Рис. 22 .2) основывается на анализе конкретных тестовых значений, которые подготавливаются специалистами путем неформального анализа переменных и предикатов в тексте программы. При каждом тесте программа исполняется по определенному маршруту, и этот реализованный маршрут вместе с соответствующим ему потоком данных рассматриваются как эталонные компоненты структуры программы. На каждом этапе отладки выделяются непроверенные компоненты программы, для которых необходимо подготавливать тесты.

Данная стратегия хороша при сравнительно простой структуре программы и преобладании в ней вычислительных операторов. Акцентируется внимание на вычислительной части программы. Однако возрастает риск пропустить сочетание предикатов, определяющих непроверенный маршрут.

При *восходящем тестировании* сначала проверяются модули нижних уровней, к которым последовательно подключаются вызывающие их модули. В этих модулях отладка также начинается с простейших конструкций, переменных и маршрутов обработки информации. Последовательное наращивание компонент программ снизу-вверх позволяет проверять работоспособность таких групп в их естественном исполнении, без подмены и имитации компонент нижних уровней. Основные трудности при такой стратегии состоят в непрерывном обновлении и увеличении числа тестовых наборов по мере подключения каждой новой компоненты более высокого уровня. Однако одновременно углубляется тестирование компонент нижних иерархических уровней. В результате может быть отлажен базовый набор модулей, пригодных для повторного использования.

При *нисходящем тестировании* отладка начинается с программ организации вычислительного процесса. Первоначально тестируется управляющее ядро комплекса программ и программы решения функциональных задач, размещенных на высших уровнях. К ним последовательно подключаются компоненты более низких иерархических уровней. Такая стратегия особенно эффективна, если имеется набор ранее проверенных программных модулей. Если некоторые программы нижних уровней не разработаны, то вместо них подключаются имитаторы “заглушки”. Преимуществом такой стратегии является сохранение и последовательное развитие тестовых исходных данных по мере подключения компонент. Однако тестирование программ с “заглушками” может требовать больших затрат на обнаружение простейших ошибок в подключаемых модулях, если они ранее не тестировались.

Этапы и задачи тестирования программных компонент.

1. *Тестирование спецификаций* состоит в проверке

- полноты и взаимного соответствия функций программ и информационных компонентов,

- соответствия описания информации на входах и выходах взаимодействующих программных модулей и групп программ,

- описания информационных модулей в базах данных.

Тестирование целесообразно проводить по нисходящему методу, начиная от спецификации комплекса или группы программ. Последовательно по уровням должно проверяться обеспечение программ верхнего уровня функциями программ нижних уровней, предписанными их спецификациями. Проверяется полнота выполнения этих функций спецификациями информационных модулей.

2. *Тестирование интерфейсов* в спецификациях программных компонент применяется для обнаружения ошибок в описаниях переменных и передачах управления при взаимодействии переменных и групп программ. Тестирование может производиться с помощью CASE средств автоматизации. Это тестирование особо важно, когда в работе принимают участие несколько исполнителей, которые могут использовать, например, разные описания переменных.

3. *Тестирование и отладка функций* в программных модулях состоит в проверке корректности обработки модулями поступающей информации и соответствия выходных данных функциям, представленным в спецификациях.

Проверке подлежат маршруты обработки информации в каждом модуле и правильность их реализации в зависимости от исходных данных. Полнота теста определяется степенью покрытия возможных маршрутов исполнения программы. На каждом выделенном маршруте должна проверяться корректность выполняемых вычислений при некоторых фиксированных исходных данных. Для каждого выделенного маршрута формируется набор условий, определяющий его реализацию. Необходимо контролировать достигнутую степень проверки маршрутов.

4. *Тестирование вычислений и преобразования данных* служит для обнаружения ошибок в вычислительной части программ. Эталонами служат результаты предварительных расчетов по формулам, заложенным в модуле. Тестирование необходимо проводить на упорядоченных наборах данных с учетом степени их влияния на выходные результаты. Первый вид наборов соответствует исходным данным в критических точках и на границах областей изменения переменных. При таких критических значениях переменных может меняться маршрут исполнения программы, вследствие чего возможно наибольшее изменение результатов. Второму виду наборов соответствуют данные в ограниченной или неограниченной области определения. Изменение данных внутри такой области не влияет на маршрут исполнения программы, поэтому при тестировании достаточно использовать только несколько значений внутри и вблизи границ области.

5. *Тестирование полноты функций*, выполняемых модулем, предназначено для выявления ошибок в выполнении функции, заданных в программной спецификации. Необходимо тестировать реализацию всех функций, определяемых спецификацией. При этом используются результаты предшествующего тестирования структуры и вычислений. Необходимо охватить все возможные варианты исходных данных, в том числе аномальные.

6. *Тестирование программных компонент* предназначено для проверки корректности управляющих и информационных связей между группами модулей, а также корректность вычисляемых функций в процессе обработки информации в группе программ. При этом значительно возрастает

сложность тестируемых объектов и, соответственно, объем и сложность тестов. В ряде случаев результаты следует получать методами стохастического тестирования.

7. *Тестирование структуры группы программ* заключается в проверке правильности вызовов модулей и возвратов управления. Реализующая структура группы может быть построена CASE средствами автоматизировано. Эта структура также автоматизировано CASE средствами может сопоставляться с формализованной спецификацией на группу программ. Различие между структурами квалифицируется либо как ошибки, либо как изменение, обусловленное развитием проекта.

8. *Тестирование выполнения ограничений по памяти и длительности исполнения*, кроме очевидной проверки, содержит проверку исполнения задач БЗУ в реальном времени.

9. *Тестирование полноты решения функциональной задачи группой программ* служит для завершающего тестирования группы программ. Предыдущие тестирования показали, что все модули исправны, их схемы вызовов проверены. Теперь необходимо проверить всю систему. При этом надо проверить логику управления системой, где большое значение имеют тесты для проверки работы в особых точках исходных данных и при критическом сочетании условий. Проверка правильности алгоритмов вычислений была проверена при проверке модулей.

Рекомендуемое оформление результатов тестирования в курсовых и дипломных проектах.

Должны быть подготовлены исходные данные для тестирования по всем девяти вышеизложенным пунктам. Данные помещаются в раздел «Анализ данных для тестирования» технологической части проекта. Отсутствие данных для какого-либо пункта аргументируется.

Результаты тестирования оформляются в виде протоколов тестирования. Протокол тестирования – это таблица из трех столбцов. Первый столбец – исходные данные для тестирования. Второй столбец – это ожидаемые заранее вычисленные или полученные путем логических размышлений результаты. Третий столбец – это полученные при тестировании результаты. Очевидно, что первый и третий столбцы должны полностью совпадать. Внизу протокол подписывается исполнителем, который с этого момента отвечает за соответствие своего продукта ТЗ.

Помните! Ваши программы индивидуальны, поэтому процесс тестирования в каждом случае индивидуален. От Вашей эрудиции, мастерства программиста, глубины освоения своего программного продукта, таланта составления диагностических тестов и описания результатов тестирования зависит Ваш успех на рынке продаж.

Список литературы

1. Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д. Математические методы в теории надежности.- М. Наука,1965.

2. Вентцель Е. С., Исследование операций. – М. Советское радио, 1972.

3. Липаев В. В., Надежность программных средств. Серия «Информатизация России на пороге XXI века». – М.: СИНТЕГ, 1998.

4 Черчесов Г. Н., Надежность программных комплексов. Учебное пособие. – СПб.: Питер, 2005.

5 Иыуду К. А., Надежность контроль и диагностика вычислительных машин и систем: Учебное пособие. – М.: Высшая школа, 1989.

СОДЕРЖАНИЕ.. Стр.