

Высшее образование

И. А. Ушаков

КУРС ТЕОРИИ НАДЕЖНОСТИ СИСТЕМ

дрофа



И. А. Ушаков

КУРС ТЕОРИИ НАДЕЖНОСТИ СИСТЕМ



Допущено
Министерством образования Российской Федерации
в качестве учебного пособия
для студентов высших учебных заведений,
обучающихся по специальности 010500 Механика

ДРОФА 
МОСКВА 2008

УДК 004(075.8)
ББК 30.14я73
У93

Рецензенты:

д-р физ.-мат. наук, проф. *В. А. Каштанов* (зав. каф. «Исследование операций» Московского государственного института электроники и математики);
д-р техн. наук, проф., засл. деят. науки РФ *С. А. Тимашев* (директор научно-инженерного центра «Надежность и ресурс больших систем машин» УрО РАН)

В оформлении обложки и титульного листа использована литография *Моурица Корнелиуса Эшера* «Cubic Space Division» («Kubische ruimteverdeling»), 1952, The Escher Museum at The Hague, The Netherlands

Ушаков, И. А.

У93 Курс теории надежности систем : учеб. пособие для вузов / И. А. Ушаков. — М. : Дрофа, 2008. — 239, [1] с. : ил.

ISBN 978-5-358-01586-9

Учебное пособие предназначено как для инженерных вузов, так и для математических факультетов университетов. Будущие инженеры получают благодаря ему необходимые теоретические знания, а будущие математики-прикладники — представление о сути инженерных задач и методах их формализации.

При работе с книгой не понадобятся специальные знания, выходящие за рамки стандартных математических курсов. Необходимые дополнительные сведения имеются в приложениях. Приводится много примеров, взятых из инженерной практики.

Для студентов, аспирантов и преподавателей технических вузов и университетов.

УДК 004(075.8)
ББК 30.14я73

Учебное издание

Ушаков Игорь Алексеевич

КУРС ТЕОРИИ НАДЕЖНОСТИ СИСТЕМ

Учебное пособие для вузов

Зав. редакцией *Б. В. Понкратов*. Ответственный редактор *Е. А. Вольмир*
Художественное оформление *Ю. В. Христинч*. Технический редактор
И. В. Грибкова. Компьютерная верстка *А. В. Маркин*
Корректор *Г. И. Мосякина*

Санитарно-эпидемиологическое заключение № 77.99.02.953.Д.006315.08.03 от 28.08.2003.

Подписано к печати 17.09.07. Формат 60×90 1/16. Бумага типографская. Гарнитура «Школьная». Печать офсетная. Усл. печ. л. 15,0. Тираж 3000 экз. Заказ № 6759.

ООО «Дрофа». 127018, Москва, Сущевский вал, 49.

По вопросам приобретения продукции издательства «Дрофа» обращаться по адресу: 127018, Москва, Сущевский вал, 49. Тел.: (495) 795-05-50, 795-05-51. Факс: (495) 795-05-52.

Торговый дом «Школьник». 109172, Москва, ул. Малые Каменщики, д. 6, стр. 1А. Тел.: (495) 911-70-24, 912-15-16, 912-45-76.

Магазины «Переpletные птицы»:

127018, Москва, ул. Октябрьская, д. 89, стр. 1. Тел.: (495) 912-45-76;
140408, Московская обл., г. Коломна, Голутвин, ул. Октябрьской революции, 366/2. Тел.: (495) 741-59-76.

Интернет-магазин: <http://www.drofa.ru>

Отпечатано с предоставленных диапозитивов
в ОАО «Тульская типография». 300600, г. Тула, пр. Ленина, 109.

ISBN 978-5-358-01586-9

© ООО «Дрофа», 2008

Посвящается памяти А. И. Станкевича

Предисловие

Написание этой книги было инициировано профессором Александром Ивановичем Станкевичем, с которым мы долгое время сотрудничали и дружили. Много лет я работал консультантом руководимого им отдела в одном из оборонных НИИ. Впоследствии во время своего визита ко мне в Сан-Диего он предложил написать для студентов технических вузов и инженеров учебник, в котором были бы изложены практические методы теории надежности.

Александр Иванович многие годы заведовал кафедрой машиноведения и деталей машин в МАИ и одновременно работал заместителем Генерального директора издательства «Дрофа», которое все хорошо знают по огромному вкладу, вносимому им в дело образования в России. В «Дрофе» Александр Иванович возглавил новое для издательства направление — создание учебной литературы для высшего и среднего специального образования.

Александр Иванович Станкевич оставил о себе память не только как замечательный учитель. Он был отзывчивым человеком, готовым помочь каждому в трудные минуты жизни. И это не дежурные слова: это его качество я испытал на себе.

Предлагаемый текст родился не за письменным столом, а в процессе инженерных и прикладных разработок автора в России (НИИ автоматической аппаратуры, НИИ систем связи и управления, ВЦ РАН) и в Соединенных Штатах Америки (SOTAS, QUALCOMM, Hughes Network Systems). Почти все материалы были представлены на лекциях как в России (Московский физико-технический институт, Московский энергетический институт), так и в США (George Washington University, George Mason University, University

of California San Diego), а также докладывались на международных конференциях в Болгарии, Венгрии, Германии, Италии, Канаде, Норвегии, Франции, Швеции, США и Японии.

В заключение мне хотелось бы выразить искреннюю признательность Александру Владимировичу Бочкову за большую помощь при подготовке рукописи к печати.

Сан-Диего
Калифорния

И. Ушаков

Глава 1

Показатели надежности

Эффективность, качество и экономичность использования технических систем характеризуются целым рядом показателей. Важнейшее место среди них занимают показатели надежности. Приводится анализ применяемых на практике показателей надежности.

§ 1.1. РОЛЬ ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ

Говоря о надежности, нужно ввести некоторый критерий качества функционирования, который позволил бы дать определение отказа системы. В силу множества причин, его порождающих, отказ системы является событием случайным, т. е. время успешной работы системы до отказа случайно. Восстановление системы после отказа (если оно в принципе возможно) также происходит за случайное время. В связи с этим для описания основных показателей надежности может быть в полной мере использован аппарат теории вероятностей и соответствующие характеристики случайных величин (см. Приложение 1).

Иными словами, можно сказать, что показатели надежности представляют собой характеристики случайных величин и процессов, сформулированные на инженерном языке.

Прежде чем начата работа над технической системой, заказчик формулирует техническое задание. Создавая систему, разработчик проводит предварительные расчеты и необходимые эксперименты, чтобы быть уверенным, что проектируемая система удовлетворит требованиям, предъявленным заказчиком. Для обеспечения надежности на ранней стадии разработчик проводит аналитические расчеты или моделирование для подтверждения правильности проектирования. Затем наступает время специальных испытаний на надежность, когда ее требуемый уровень подтверждается экспериментально. Наиболее же полная проверка способности системы безотказно функционировать осуществляется уже на этапе ее опытной эксплуатации.

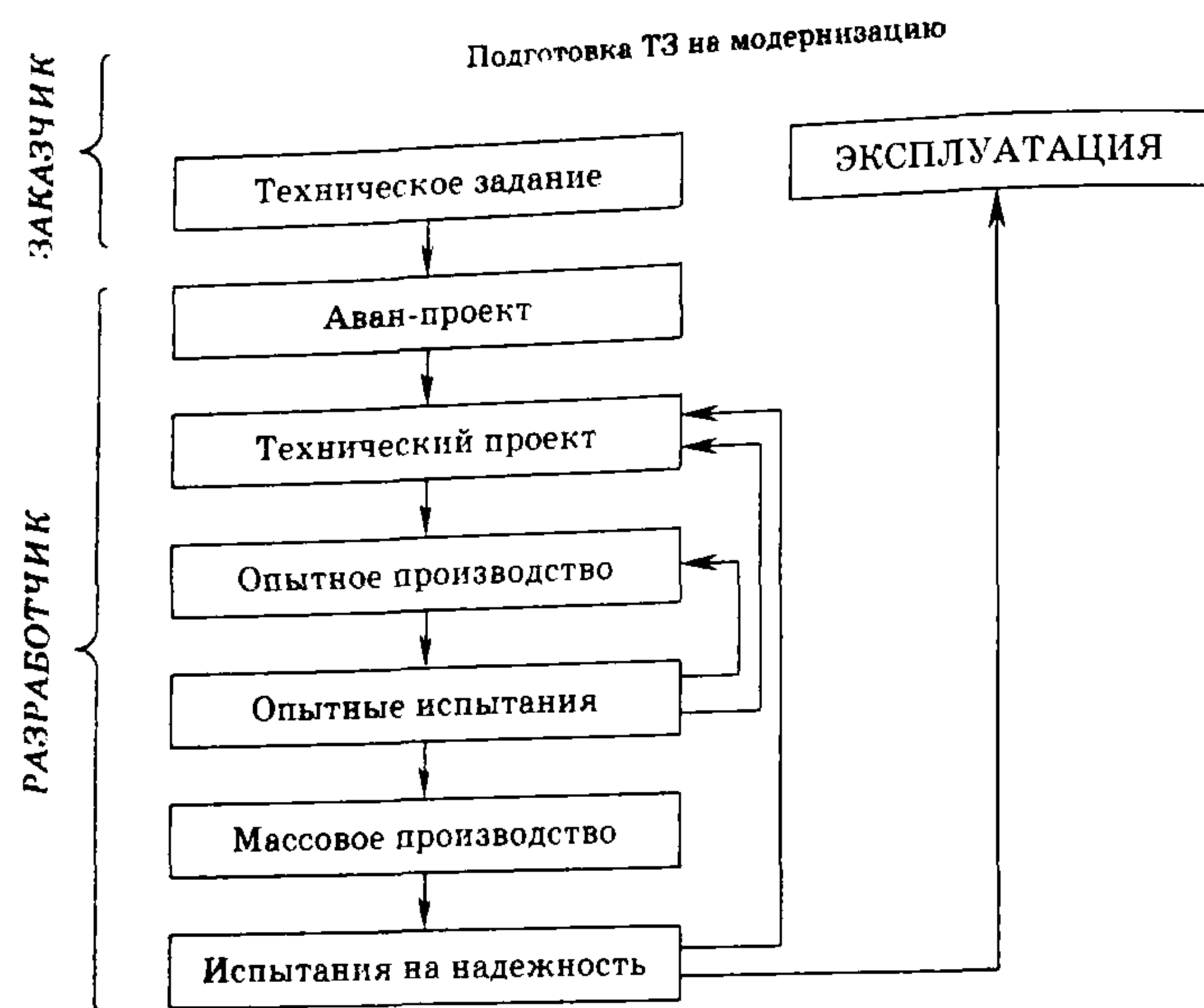


Рис. 1.1

Таким образом, «жизненный цикл» системы можно представить в виде блок-диаграммы, представленной на рис. 1.1. На этой блок-диаграмме стрелки, направленные снизу вверх, обозначают обратную связь: результаты, полученные на последующих этапах, анализируются и используются на предшествующих этапах (вносятся корректировки в структуру системы, меняется номенклатура комплектующих элементов и т. п.).

Подобная обратная связь способствует «естественному отбору» систем в процессе технической эволюции: те системы, которые работают эффективнее и/или требуют меньших затрат на производство и эксплуатацию, «выживают», и далее, за счет модернизации — «направленных мутаций» — порождают следующее поколение систем.

§ 1.2. НЕВОССТАНАВЛИВАЕМЫЕ ОБЪЕКТЫ

Определение невосстанавливаемого объекта

Заметим сразу, что деление объектов на восстанавливаемые и невосстанавливаемые является весьма условным. Например, является ли компьютер восстанавливаемым объектом? Если он ис-

пользуется для повседневной работы с текстами или несложных расчетов, то ответ — да. Однако если тот же компьютер используется для управления длительным непрерывным технологическим процессом, то восстановление его работоспособности в случае отказа приведет к недопустимой приостановке этого процесса, т. е. к отказу системы в целом. Действительно, то, что компьютер может быть восстановлен впоследствии, не играет никакой роли при рассмотрении данного непрерывного технологического процесса. Если же в системе предусмотрен активный резервный компьютер, подхватывающий функции отказавшего компьютера без ущерба для технологического процесса, то такая система, состоящая из двух компьютеров, уже может рассматриваться как восстанавливаемая: ее работоспособность может быть восстановлена, хотя и с возможным перерывом в работе.

1.2.2. Критерии отказов

Поведение невосстанавливаемого объекта описывается случайной величиной — временем до отказа. Говоря о надежности того или иного технического объекта, следует всегда предполагать, что отказ этого объекта четко определен. Без сформулированного понятия *отказ* невозможно правильно истолковать, что такое *время до отказа*, а следовательно, задать и показатели надежности.

Отказом может быть либо полное и мгновенное прекращение функционирования системы, либо снижение эффективности ее функционирования до некоторого недопустимого (заданного) уровня (рис. 1.2). Обычно отказы по этим признакам подразделяются на два типа: внезапные и постепенные.

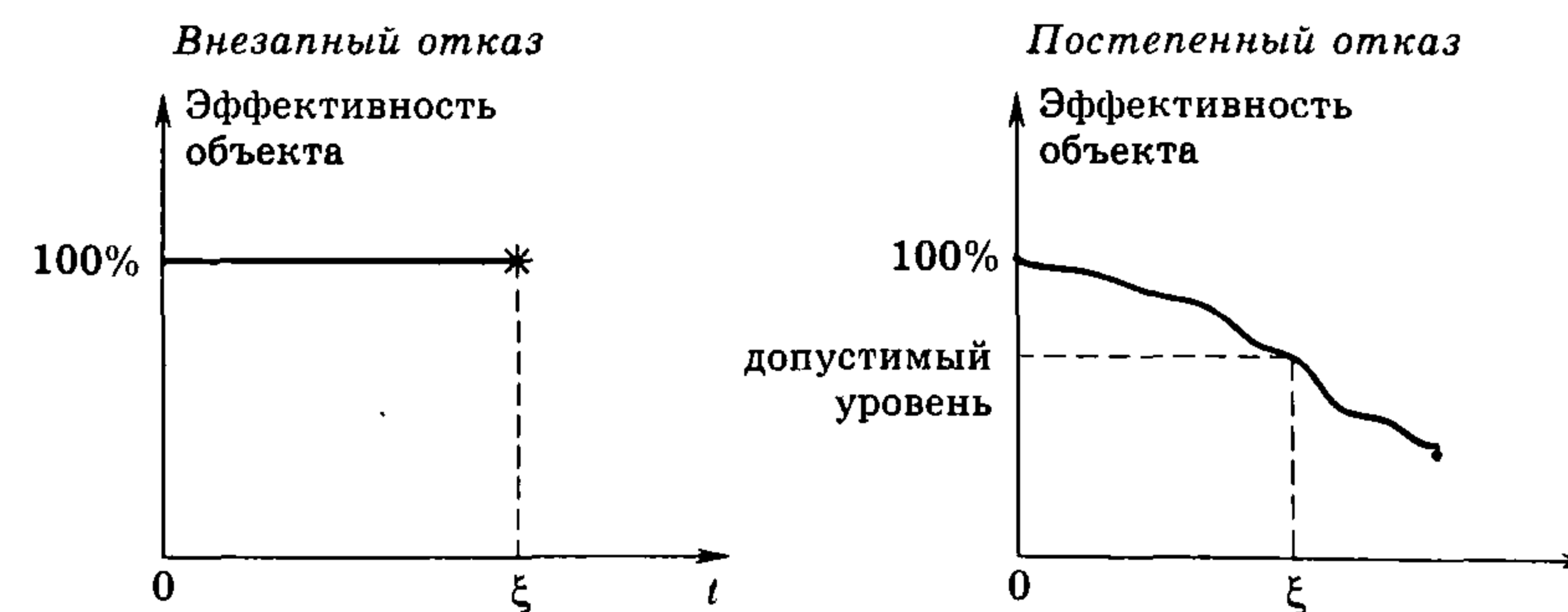


Рис. 1.2

Часто технические объекты работают не непрерывно, а включаются лишь на время выполнения тех или иных операций. В этих случаях рассматривается только *чистая* наработка объекта, т. е. интервалы отключения объекта не учитываются. Таким образом, ось времени *спрессовывается* и процесс функционирования объекта представляет собой как бы непрерывное его включение в работу.

1.2.3. Вероятность безотказной работы

Обозначим случайное время работы объекта до отказа через ξ .

Вероятность безотказной работы (ВБР) объекта за заданное время t_0 определяется как

$$P\{\xi > t_0\} = 1 - F(t_0), \quad (1.1)$$

где $P(t)$ есть функция распределения случайной величины ξ .

Если на основании специальных испытаний или эксплуатации известны N наблюдаемых значений (реализаций) $\xi_1, \xi_2, \dots, \xi_N$ случайной величины ξ , то статистическая оценка вероятности безотказной работы имеет вид (рис. 1.3)

$$\hat{P}(t_0) = \frac{1}{N} \sum_{k=1}^N d_k(t_0), \quad (1.2)$$

$$d_k(t_0) = \begin{cases} 1, & \text{если } \xi_k \geq t_0, \\ 0, & \text{если } \xi_k < t_0, \end{cases}$$

т. е. реализации приписывается значение 1, если объект успешно проработал до момента t_0 , и значение 0 в противном случае. Полученная оценка является статистически несмещенной. Она показывает,

Наработка

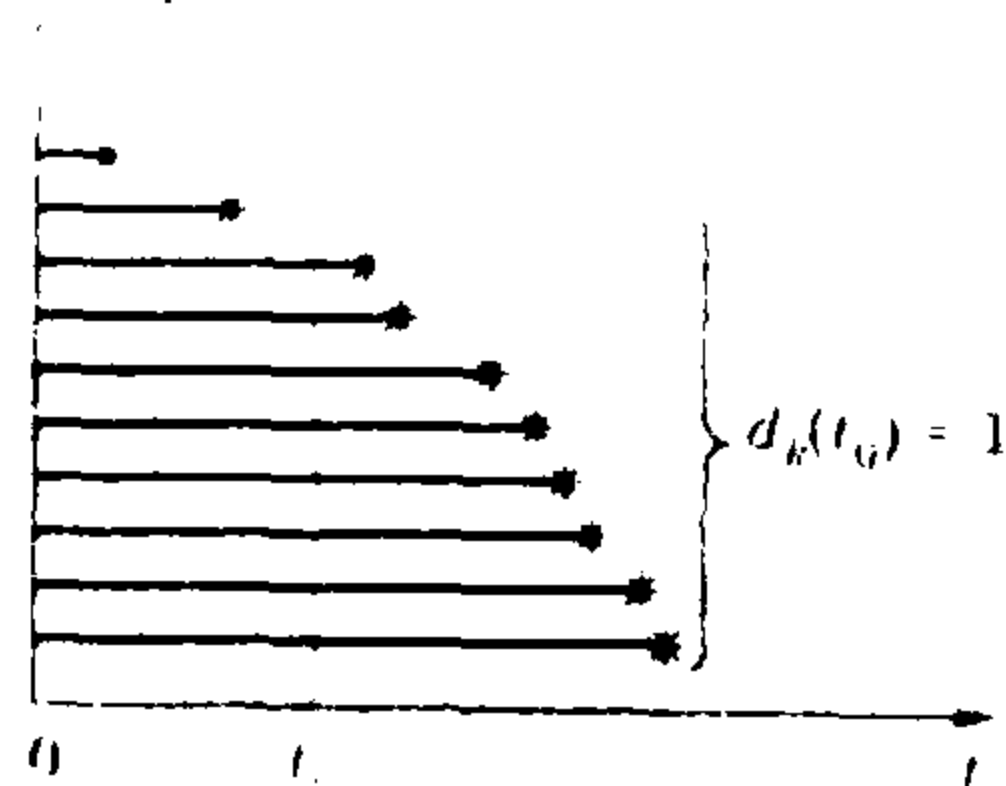


Рис. 1.3

какая доля испытанных объектов проработала заданное время t_0 .

В теории надежности часто используется экспоненциальное распределение* времени безотказной работы. В этом случае выражение для ВБР принимает следующий вид:

$$P(t_0) = \exp\left(-\frac{t_0}{T}\right), \quad (1.3)$$

здесь T есть среднее время безотказной работы

* в теории надежности распределение (Приложение 2).

Приложение

Само время t_0 может быть случайной величиной с распределением $G(t)$, т. е. время выполнения объектом своей операции не постоянно. В этом случае ВБР объекта вычисляется по формуле

$$P = \int_0^{\infty} P(t) dG(t) = \int_0^{\infty} P(t) g(t) dt, \quad (1.4)$$

в которой $g(t)$ — плотность распределения $G(t)$.

Такая постановка задачи совершенно естественна для многих систем: вычислительные задачи имеют различную трудность, транспортные системы бывают различного времени в пути, и т. п.

ПРИМЕР 1.1. Пусть система характеризуется экспоненциальным распределением времени работы до отказа, $F(t) = 1 - \exp(-\lambda t)$, а время выполнения системой заданной операции — случайная величина, имеющая произвольное распределение $G(t)$ со средним θ и дисперсией σ^2 . Допустим, что $\lambda\theta \ll 1$. Найти приближенное решение для вероятности безотказной работы

Решение. В данном случае можно записать приближение

$$P(t) = e^{-\lambda t} \approx 1 - \lambda t + 0,5(\lambda t)^2.$$

Следовательно, решение приобретает вид

$$\begin{aligned} P &\approx \int_0^{\infty} [1 - \lambda t + 0,5(\lambda t)^2] dG(t) = \\ &= \int_0^{\infty} dG(t) - \lambda \int_0^{\infty} t dG(t) + 0,5\lambda^2 \int_0^{\infty} t^2 dG(t) = 1 - \lambda\theta + 0,5\lambda^2 M_2, \end{aligned} \quad (1.5)$$

где M_2 — второй начальный момент распределения $G(t)$.

Из выражения (1.5) следует, что замена случайного времени выполнения задачи на среднее привела бы к занижению истинного значения ВБР.

1.2.4. Среднее время работы до отказа

Одним из самых простых показателей надежности невосстанавливаемых объектов является среднее время работы до отказа, т. е. математическое ожидание случайной величины ξ .

Среднее время работы до отказа определяется как

$$T = E\{\xi\} = \int_0^{\infty} x f(x) dx = \int_0^{\infty} x dF(x), \quad (1.6)$$

где $f(x)$ и $F(x)$ — соответственно плотность и функция распределения случайной величины времени работы до отказа

* Интегрирование ведется от нуля, так как время работы есть неотрицательная случайная величина.

Для обозначения математического ожидания случайной величины ξ используется символ $E\{\xi\}$.

Если известны экспериментальные реализации $\xi_1, \xi_2, \dots, \xi_N$ случайной величины ξ , то статистическая оценка среднего времени работы до отказа вычисляется как

$$T = \frac{1}{N} \sum_{k=1}^N \xi_k. \quad (1.7)$$

Полученная оценка является статистически несмещенной.

1.2.5. Интенсивность отказов

Для характеристики изменения надежности во времени удобно использовать функцию интенсивности отказов, которая обычно обозначается через $\lambda(t)$. В теории надежности эта функция имеет тот же смысл, что и «кривая смертности» в демографии, и определяется аналогичным образом с точностью до наименований.

Формально функция интенсивности отказов определяется как

$$\lambda(t) = \frac{f(t)}{P(t)}, \quad (1.8)$$

т. е. $\lambda(t)$ есть условная плотность вероятности отказа в момент t , если до этого момента отказ еще не произошел.

Типичная кривая смертности $\lambda(t)$ представлена на рис. 1.4.

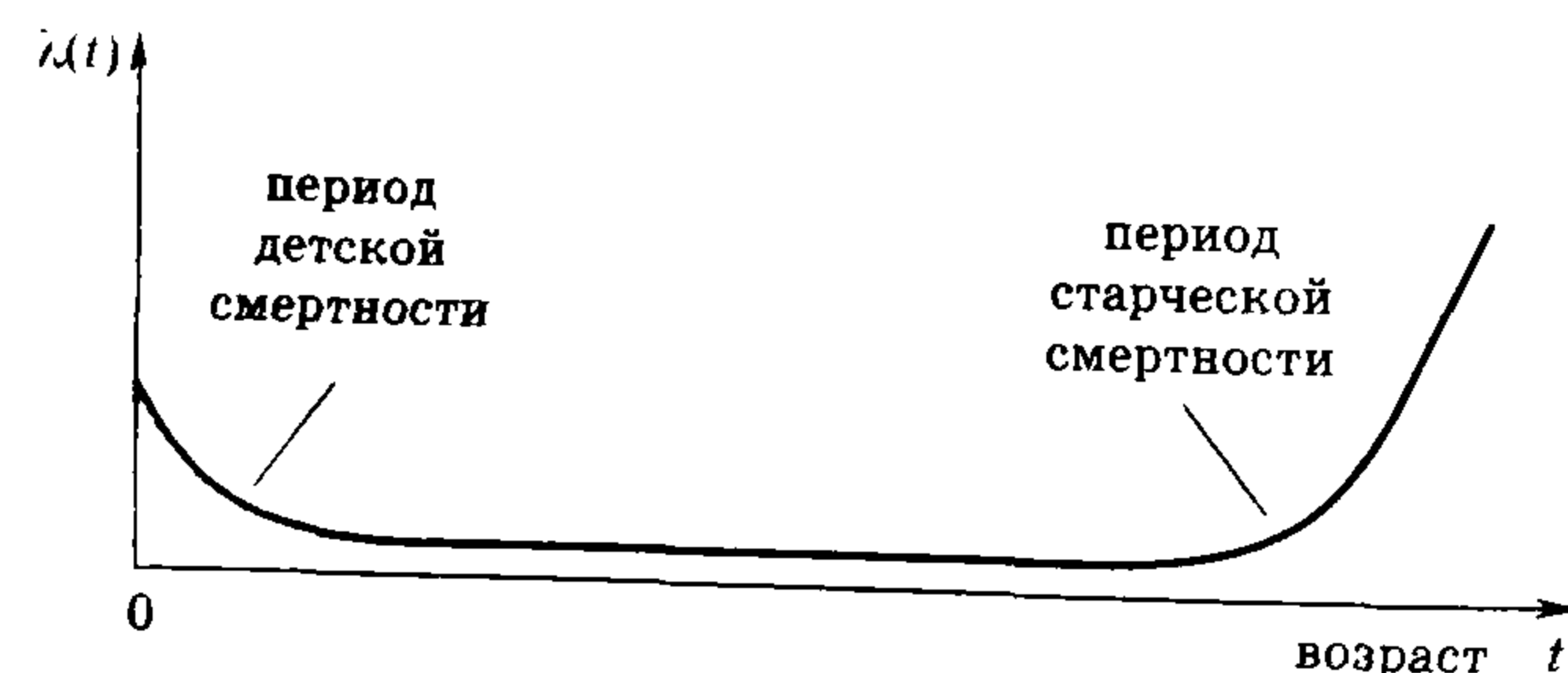


Рис. 1.4

Физический смысл функции $\lambda(t)$ становится понятным, если записать эквивалентное выражение

$$\lambda(t) \Delta t = \frac{f(t) \Delta t}{P(t)}, \quad (1.9)$$

в котором

- ✓ величина $f(t) \Delta t$ — абсолютная вероятность отказа во временном интервале $[t, t + \Delta t]$;
- ✓ $\lambda(t) \Delta t$ — условная вероятность отказа объекта в интервале $[t, t + \Delta t]$ при условии, что объект не отказал до момента t ;
- ✓ $P(t)$ — вероятность того, что объект не отказал до момента t .

Таким образом, (1.9) представляет собой выражение для условной вероятности.

Формулу (1.8) можно переписать в виде

$$\lambda(t) = \frac{1}{P(t)} \cdot \frac{d}{dt} F(t) = -\frac{1}{P(t)} \cdot \frac{d}{dt} P(t) = -\frac{d[\ln P(t)]}{dt}. \quad (1.10)$$

Отсюда следует, что для любой функции распределения можно записать:

$$P(t) = \exp\left(-\int_0^t \lambda(x) dx\right). \quad (1.11)$$

Экспоненциальная форма записи (1.11) иногда трактуется следующим образом: «любое распределение является экспоненциальным». Безусловно, это не так, поскольку экспоненциальным называется распределение, у которого показатель есть *линейная функция аргумента*.

Из фундаментального условия $0 \leq P(t) \leq 1$ следует, что

$$0 \leq \int_0^t \lambda(x) dx < \infty, \quad (1.12)$$

$$\lim_{t \rightarrow \infty} \int_0^t \lambda(x) dx = \infty. \quad (1.13)$$

Теоретически предел в (1.13) может быть конечным, что соответствует несобственному распределению с бесконечным математическим ожиданием. Здесь случайная величина есть время работы до отказа, т. е. не может быть бесконечной по своей физической сущности.

Функция интенсивности отказов хорошо описывает физические процессы, происходящие с реальными объектами. Основные типы распределений делятся на

- ✓ *стареющие*, у которых функция распределения монотонно возрастающая, — к ним относятся такие часто встречающиеся на практике распределения, как нормальное и эрланговское;
- ✓ *молодеющие*, у которых функция распределения монотонно убывающая, — к ним относятся, в частности, взвешенное экспоненциальное распределение.

Экспоненциальное распределение является граничным для обоих упомянутых выше классов: функция интенсивности для него есть постоянная λ , равная $1/T$, где T — среднее время работы до отказа.

Особое место занимает распределение* Вейбулла — Гнеденко, описывающее весьма широкий класс распределений неотрицательных случайных величин, как *стареющих*, так и *молодеющих*.

Замечание. Экспоненциальное распределение обладает марковским свойством. Это означает, что интенсивность отказов не зависит от того, сколько уже проработал данный объект. Конечно, такое допущение является в общем-то достаточно сильным и должно приниматься с определенными оговорками. Используя аналогию с демографией, можно сказать, что экспоненциальному распределению подчиняется смертность от «падения на голову кирпича с крыши», т. е. ситуации, когда причина смерти не зависит от возраста. Для экспоненциального распределения среднее время до отказа T является величиной, обратной интенсивности отказов λ . Здесь нужно отметить, что реальный срок жизни технических объектов в силу естественного старения может оказаться меньше среднего времени до отказа!

ПРИМЕР 1.2. Поясним сказанное выше с помощью демографического примера. Представим себе, что в некотором спортивном лагере собрано 10 000 здоровых молодых людей. Каждый год производится смена состава. Смерть молодого человека возможна лишь по чисто случайной и непредвиденной причине. (Допустим, что пьянство, драки и прочие опасные для здоровья факторы в спортлагере отсутствуют.) Думаю, читатель согласится: смерть более чем двух-трех человек за год будет считаться ненормально высокой. Это означает, что «интенсивность отказов» в спортивном лагере в среднем заведомо меньше чем $3 \cdot 10^{-4}$ [1/год]. Исходя из такой величины интенсивности, определяем, что среднее время «безотказной работы» человека в спортивном лагере более 3000 лет! Но ведь смерть человека за счет естественного старения организма наступает заведомо до того, когда он достигнет 150-летнего возраста.

Иными словами, в нашем примере среднее время *безотказной работы* в 20 раз превышает среднюю длительность жизни. Это среднее время характеризует на самом деле не человека, как такового, а *спальное место*, на котором, сменяясь, спят молодые и здоровые люди!

§ 1.3. ВОССТАНАВЛИВАЕМЫЕ ОБЪЕКТЫ

1.3.1. Определение восстанавливаемого объекта

Как уже отмечалось, деление всех объектов на два класса — *невосстанавливаемых* и *восстанавливаемых*, является достаточно условным.

* См. Приложение 2, с. 215.

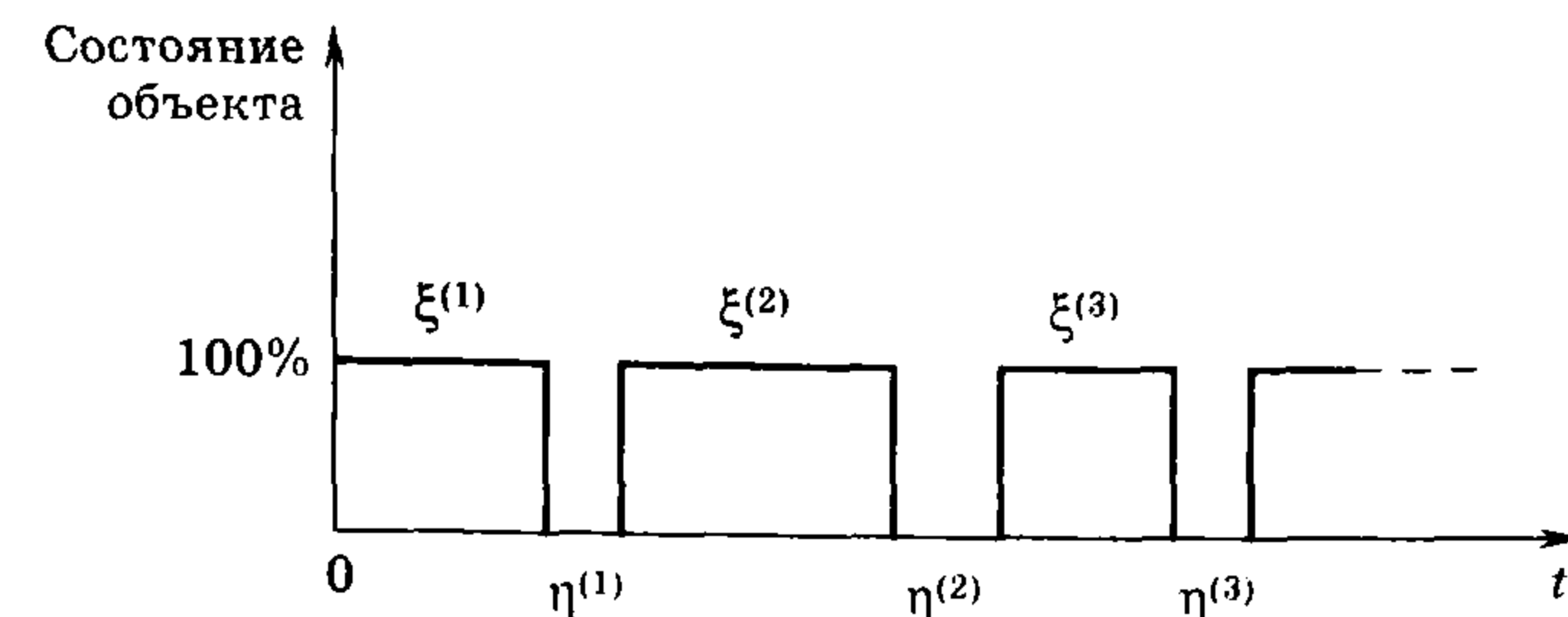


Рис. 1.5

Если объект предназначен для длительного использования и его отказ либо задерживает выполнение операции, либо даже приводит к отказу части операций, но функционирование объекта может быть возобновлено после проведения восстановительных операций, то говорят о *восстанавливаемых объектах*.

Поведение восстанавливаемого объекта описывается случайным альтернирующим точечным процессом, представляющим собой чередование периодов успешного функционирования ξ и периодов восстановления η , представленных на рис. 1.5, где введены следующие обозначения: $\xi^{(k)}$ — интервал времени до k -го отказа после предыдущего восстановления; $\eta^{(k)}$ — длительность k -го восстановления.

Обычно объекты с восстановлением работают в непрерывном режиме, т. е. без периодов простаивания вхолостую. Даже такие объекты, которые выполняют свои функции от задачи к задаче, в интервалах между выполнением задач должны находиться в режиме готовности (например, системы управления воздушным движением в аэропортах, системы противовоздушной обороны и т. п.).

Если невосстанавливаемые объекты могут быть самых разных физических размеров — от микросхем до систем, то восстанавливаемые объекты — это обычно большие технические системы. В дальнейшем будем использовать термин *система*, поскольку он конкретнее отражает специфику восстанавливаемого объекта.

1.3.2. Среднее время работы между отказами

Время работы до первого отказа, $\xi^{(1)}$, вообще говоря, отличается от времени работы до второго отказа, $\xi^{(2)}$, оба отличаются от времени работы до третьего отказа, $\xi^{(3)}$, и т. д. Это связано с тем, что система состоит из многих элементов, ее структура может включать избыточные элементы, а само восстановление может выполняться по различным правилам.

ПРИМЕР 1.3. Рассмотрим дублированную систему из двух независимых элементов. В начальный момент времени работоспособны оба ее элемента. Начав работать в этот момент времени, система работает некоторое среднее время до первого отказа, $T^{(2)}$. После отказа системы, который наступает, когда оба элемента отказали, происходит ее восстановление. В некоторый случайный момент времени один из элементов восстанавливается и система опять начинает функционировать. Однако на этот раз начальным будет состояние с единственным работоспособным элементом, пример возможной реализации которого представлен на рис. 1.6. Обозначим среднее время работы до отказа для этих начальных условий через $T^{(1)}$.

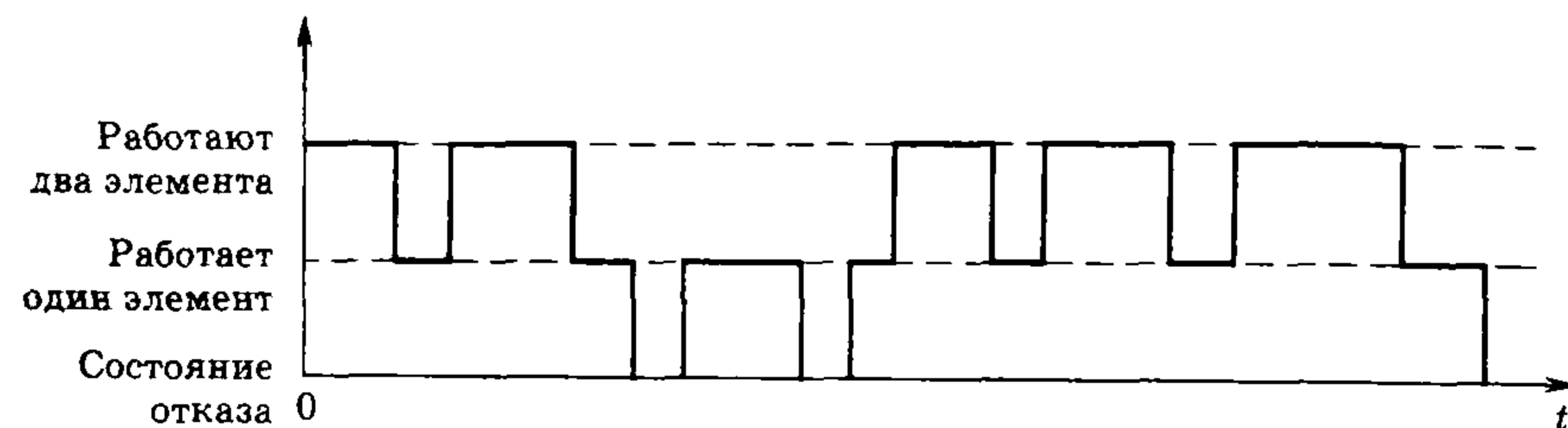


Рис. 1.6

По рис. 1.6 видно, что среднее время работы системы до первого отказа при исходной работоспособности обоих этих элементов равно

$$T^{(2)} = T^* + T^{(1)};$$

здесь T^* — время до появления первого отказа любого из элементов.

Примерно такого же рода рассуждения могут быть повторены и в более общих случаях.

Большинство систем с восстановлением функционирует многие тысячи часов, за это время происходит достаточно большое число отказов, поэтому на практике не делают различий между средними различными величинами $\xi^{(1)}$, $\xi^{(2)}$, Исключение составляет, быть может, только среднее время до первого отказа.

Обычно говорят о среднем времени работы между отказами T , понимая его как среднее время между отказами в стационарном (установившемся) процессе функционирования системы.

Среднее время работы между отказами определяется аналогично выражению для среднего времени работы до отказа (1.6):

$$T = E\{\xi\} = \int_0^{\infty} x f(x) dx = \int_0^{\infty} x dF(x); \quad (1.14)$$

здесь $f(x)$ и $F(x)$ — соответственно плотность и функция распределения случайной величины ξ — времени работы между отказами.

На практике, если наблюдаются n систем и для них известны значения N_1 реализаций для первой системы $\xi_1^{(1)}, \xi_2^{(1)}, \dots, \xi_{N_1}^{(1)}$, N_2 реализаций для второй $\xi_1^{(2)}, \xi_2^{(2)}, \dots, \xi_{N_2}^{(2)}$, ... и N_n реализаций для n -й системы $\xi_1^{(n)}, \xi_2^{(n)}, \dots, \xi_{N_n}^{(n)}$, то среднее время между отказами вычисляется как

$$T = \frac{1}{N_1 + N_2 + \dots + N_n} \sum_{k=1}^n \sum_{j=1}^{N_k} \xi_j^{(k)}. \quad (1.15)$$

1.3.3. Среднее время восстановления

Время восстановления является фактически интервалом, в течение которого система не может продолжать выполнение своих функций. Для одних систем это может быть время собственно ремонта отказавшей части, для других — время замены отказавшего блока или иного сменного элемента на исправный из числа запасных, для третьих — время переключения на встроенный резерв.

Если известно распределение времени восстановления $G(t) = P\{\eta \leq t\}$, то среднее время восстановления t вычисляется по обычной формуле для математического ожидания

$$t = E\{\eta\} = \int_0^{\infty} x g(x) dx = \int_0^{\infty} x dG(x), \quad (1.16)$$

где $g(x)$ — плотность распределения $G(x)$.

1.3.4. Коэффициент готовности

Нестационарный коэффициент готовности $K(t)$ есть вероятность того, что система находится в работоспособном состоянии в момент времени t . Обычно предполагается, что в момент $t = 0$ система абсолютно исправна, т. е. $K(0) = 1$, а прогноз состояния системы в момент t характеризуется коэффициентом готовности. При $t > 0$ имеем $K(t) < 1$. С течением времени в силу случайности величин ξ и η процесс устанавливается и $K(t)$ стремится к некоторой постоянной величине K , которая называется стационарным коэффициентом готовности, часто именуемым коэффициентом готовности (КГ):

$$K = \lim_{t \rightarrow \infty} K(t). \quad (1.17)$$

Нестационарный КГ для систем более или менее сложных в общем случае имеет довольно громоздкое выражение, в которое

входят распределения времени работы между отказами и времени восстановления. Этот показатель крайне редко используется на практике.

Стационарный КГ характеризует систему в *бесконечно удаленный от начала отсчета момент времени*, т. е. когда процесс переходов из состояния работоспособности в состояние восстановления становится стационарным. В этом случае среднее по времени совпадает со средним по ансамблю, иначе говоря, несущественно, как усреднять: либо рассматривая многократное испытание одного образца, либо одноразовые испытания многих образцов. Если известны средние значения времени безотказной работы и времени восстановления T и τ , то КГ можно выразить как

$$K = \frac{T}{T + \tau}. \quad (1.18)$$

Коэффициент готовности является наиболее существенным показателем надежности систем с восстановлением. Он может быть определен и на основании результатов испытаний или эксплуатации как

$$\hat{K} = \frac{\hat{T}}{\hat{T} + \hat{\tau}}, \quad (1.19)$$

т. е. при подстановке в качестве величин T и τ их статистических оценок.

1.3.5. Коэффициент интервальной готовности

Одной готовности совершить некую операцию, вообще говоря, еще мало: нужно не только, чтобы система была работоспособна в требуемый момент t , но и была бы еще в состоянии безотказно проработать в течение требуемого интервала времени t_0 .

Вероятность того, что система окажется работоспособной в некоторый момент времени t и проработает безотказно в течение интервала длиной t_0 , называется коэффициентом интервальной готовности $K^*(t, t + t_0)$.

На практике обычно рассматривается только стационарный коэффициент интервальной готовности, т. е. случай, когда $t \rightarrow \infty$:

$$K^*(t_0) = KP^*(t_0) = \frac{T}{T + \tau} P^*(t_0); \quad (1.20)$$

здесь через $P^*(t_0)$ обозначена вероятность того, что система проработает безотказно начиная с произвольного бесконечно удаленного момента времени.

Замечание. Следует иметь в виду, что в общем случае вероятность $P^*(t_0)$ не эквивалентна вероятности безотказной работы $P(t_0)$. Если $P(t_0)$ есть вероятность проработать безотказно на интервале $[0, t]$, начиная с состояния полной работоспособности, то $P^*(t_0)$ есть вероятность проработать в течение интервала длительностью t_0 , начиная с одного из возможных *промежуточных* состояний работоспособности. Иначе говоря, здесь мы имеем дело с *остаточным временем жизни*.

§ 1.4. ВЫБОР ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ

Надежность функционирования технических объектов характеризуется специальными показателями надежности. Эти показатели выбираются для каждого конкретного случая и должны обладать следующими необходимыми свойствами:

- ✓ отражать специфику функционирования объекта;
- ✓ иметь простой и понятный физический смысл;
- ✓ быть вычисляемыми аналитически или численно на этапе проектирования;
- ✓ быть экспериментально проверяемыми на этапах испытаний и эксплуатации.

Число показателей надежности, выбранных для характеристики объекта, должно быть по возможности небольшим (практически не превышать двух-трех). Больше число показателей лишь запутает картину. В то же время не следует прибегать к помощи *взвешенных* показателей, так как субъективно выбранные веса исказят объективную характеристику объекта.

Для различных технических объектов показатели надежности могут быть различными. Так, показатели надежности системы в первую очередь определяют надежность выполнения системой в целом требуемых операций. Отдельные подсистемы обычно характеризуются некоторыми вспомогательными показателями, а комплектующие элементы — обычно простейшими показателями надежности.

ПРИМЕР 1.4. Наземную станцию спутниковой системы сбора информации удобно характеризовать коэффициентом готовности, так как объем собранной информации пропорционален полезному времени работы.

ПРИМЕР 1.5. Если имеются две системы сбора информации с одинаковыми коэффициентами готовности, то, видимо, предпочтительной будет та из них, которая при прочих равных условиях имеет большее время работы между отказами. Это соображение следует из того, что любой отказ обесценивает целый блок данных, уже полученных до того, поэтому частые отказы приведут к большему числу потерянных блоков данных.

Надежность отдельных подсистем влияет на систему в целом через структуру системы. Показатели надежности подсистем обычно выступают в качестве расчетных параметров при расчете надежности системы в целом. Обычно для подсистем удобно выбирать два базовых показателя: среднее время работы до отказа и среднее время восстановления работоспособности. Эти показатели помогут оценить влияние надежности подсистем с учетом длительности восстановления работоспособности, выбранного правила технического обслуживания, частоты проведения профилактических работ и прочих факторов.

Нижний уровень в иерархии объектов исследования представляют неделимые, или *атомарные*, элементы. Обычно такие элементы сами по себе не являются восстанавливаемыми: при восстановлении работоспособности объекта они заменяются на новые. Такие атомарные элементы характеризуются средним временем работы до отказа (или же обратной величиной — интенсивностью отказов). Элементы сами по себе не характеризуются временем восстановления, так как оно зависит от конкретных особенностей аппаратуры: от легкости доступа к отказавшему элементу в случае ремонта или от скорости замены блока (модуля) с отказавшим элементом на запасной.

Функционирование технического объекта протекает во времени, а отказы носят случайный характер. Поэтому практически все показатели надежности имеют вероятностно-временной характер. Даже такой «мгновенный» показатель надежности, как коэффициент готовности несет временную нагрузку: это средняя доля времени, в течение которого объект является работоспособным.

§ 1.5. ЗАДАНИЕ ТРЕБОВАНИЙ ПО НАДЕЖНОСТИ

1.5.1. Иерархические уровни при задании требований

При задании требований по надежности следует различать технические объекты трех уровней.

- **Системы**, т. е. технические объекты, выполняющие самостоятельные операции.

Системы характеризуются оперативно-техническими показателями надежности и эффективности функционирования.

Примером системы может служить вычислительный комплекс для управления динамическими объектами, характеризуемый

числом одновременно обслуживаемых объектов, вероятностью безотказного управления отдельно взятым объектом и т. п. Другой пример: грузовой автомобиль с такими характеристиками, как средний пробег между отказами, вероятность выполнения стандартного рейса и т. п.

Результат работы системы может быть сформулирован в терминах выполнения требуемой или заданной операции. Показатели надежности систем выражаются через показатели надежности подсистем, входящих в их состав.

- **Подсистемы**, т. е. технические объекты, входящие в состав систем и выполняющие частные функциональные задачи.

Подсистемы, как правило, не выполняют самостоятельных завершенных операций, а служат для обеспечения выполнения операции системой в целом. В предположении выполнения подсистемами своих функциональных обязанностей они характеризуются такими показателями, как распределение времени работы до отказа и распределение времени восстановления.

- **Элементы**, т. е. базовые компоненты, являющиеся неделимыми техническими объектами (по крайней мере, в рамках текущего производимого анализа).

Элементы характеризуются средней наработкой до отказа.

Нужно иметь в виду, что подразделение технических объектов на указанные три иерархических уровня является условным: объект, который в одном случае является системой, в другом может оказаться подсистемой и т. д. Возможно иерархическое разбиение каждого иерархического уровня на несколько вспомогательных уровней. С точки зрения задания требований по надежности, как правило, оказывается достаточным иметь указанные три категории технических объектов.

1.5.2. Задание требований на систему

Задание требований на систему может осуществляться несколькими способами.

- **Экспертный метод.** Этот метод основывается на инженерной интуиции и практическом опыте, а потому не нуждается в специальных комментариях. Однако следует подчеркнуть, что экспертный метод не имеет ничего общего со «среднепотолочным» гаданием.

- **Использование аналогов или прототипов.** Метод основывается на анализе имеющейся статистической информации об эксплуатации подобных или близких по назначению систем. При этом может учитываться относительная сложность рассматриваемой системы по сравнению со сравниваемой системой, различие в условиях функционирования и т. д. Такой прогноз также в значительной степени опирается на экспертные оценки.
- **Задание оптимальных требований.** Такой подход ограничен следующими факторами:

- ✓ выходной эффект от функционирования системы измерим в тех же (обычно стоимостных) единицах, что и затраты на ее создание;
- ✓ данные о надежности элементной базы достаточно достоверны;
- ✓ полностью определены структура и принцип функционирования системы.

В случае выполнения этих условий задание требований по надежности сводится к максимизации целевой функции $\Phi(R)$ вида

$$\Phi(R) = E(R) - C(R),$$

где R — показатель надежности системы, $E(R)$ — выходной эффект от функционирования системы, $C(R)$ — затраты на обеспечение уровня надежности R .

Оптимальное значение надежности $R_{\text{опт}}$ находится из решения уравнения

$$\frac{\partial}{\partial R} E(R) = \frac{\partial}{\partial R} C(R). \quad (1.21)$$

Значение целевой функции находится как $\Phi(R_{\text{опт}})$.

Если имеется несколько возможных вариантов исполнения системы (например, несколько различных структур или возможность выбора альтернативных элементов), то решение (1.21) проводится для каждого из вариантов, получаются несколько величин $\Phi_k(R_{\text{опт}})$ и в результате выбирается наибольшее из них, что и соответствует оптимальному выбору системы.

1.5.3. Задание требований на подсистему

Предполагается, что требование по надежности на систему в целом $R_{\text{треб}}$ уже задано. Задача состоит в распределении требований между отдельными подсистемами таким образом, чтобы были удовлет-

ворены требования ко всей системе в целом. Рассматривается показатель надежности типа коэффициента готовности или вероятности безотказной работы.

- **Метод равномерного распределения квот надежности.** Если система состоит из N подсистем примерно эквивалентного объема (т. е. подсистемы близки по сложности, например, по числу элементов), то грубое значение требуемого уровня показателя надежности подсистемы k может быть найдено из условия

$$R_k = \sqrt[N]{R_{\text{треб}}}. \quad (1.22)$$

- **Метод пропорционального распределения.** Этот метод применим к системам с простой структурой (без избыточности). Если известно, что подсистема k включает в свой состав n_k некоторых эквивалентных, или приведенных, элементов, то требуемый уровень показателя надежности подсистемы k может быть найден из условия

$$R_k = (R_{\text{треб}})^{A_k}, \quad (1.23)$$

в котором A_k определяется как вес k -й подсистемы по формуле

$$A_k = \frac{n_k}{\sum_{j=1}^N n_j}. \quad (1.24)$$

Очевидно, $\sum A_k = 1$, а следовательно, $\prod R_k = R_{\text{треб}}$.

Если состав подсистем известен с большей подробностью, например, известно, что k -я подсистема включает в свой состав n_{ki} элементов типа i с интенсивностью отказов λ_i каждый, то можно ввести некий показатель типа *ожидаемого числа отказов* для каждой подсистемы:

$$\Lambda_k = \sum_{i=1}^N n_{ki} \lambda_i.$$

Требуемый уровень показателя надежности подсистемы k может быть найден из условия

$$R_k = (R_{\text{треб}})^{B_k}, \quad (1.25)$$

где B_k определяется как вес k -й подсистемы по формуле

$$B_k = \frac{\Lambda_k}{\sum_{j=1}^N \Lambda_j}. \quad (1.26)$$

Очевидно, $\sum B_k = 1$, а следовательно, $\prod R_k = R_{\text{треб}}$.

Предполагается, что подсистемы сами по себе таковы, что равное повышение надежности каждой из них требует примерно равных затрат.

- **Метод оптимального распределения квот надежности.** Если известна структура системы S , т. е. как подсистемы влияют на надежность системы в целом, а также известна зависимость надежности каждой подсистемы k от ее стоимости $R_k(C_k)$, то оптимальное распределение квот надежности находится из решения задачи:

$$\min \{ C[S, R_k(C_k)] \mid R_{\text{треб}} \}. \quad (1.27)$$

При известной структуре системы и известных зависимостях повышения надежности подсистем в зависимости от затрат задача на условную оптимизацию (1.27) решается обычными методами дискретного программирования.

ЗАДАНИЕ ТРЕБОВАНИЙ НА ЭЛЕМЕНТ выражается в следующем: если элемент является действительно атомарным с технологической точки зрения (например, микросхема или радиодеталь), то фактически возможны лишь экспертные методы задания их надежности с учетом анализа прототипов и общих тенденций развития технологии.

Глава 2

Невосстанавливаемые системы

Рассматриваются основные классы невосстанавливаемых систем. Дается описание их структурных функций. Приводятся методы расчета вероятности безотказной работы и среднего времени работы до отказа.

§ 2.1. СТРУКТУРНАЯ ФУНКЦИЯ

Состояние элемента в момент времени t удобно описывать индикаторной функцией $x_i(t)$:

$$x_i(t) = \begin{cases} 1, & \text{если элемент в момент времени } t \text{ работоспособен,} \\ 0, & \text{если элемент в момент времени } t \text{ неработоспособен.} \end{cases}$$

Очевидно, что функция $x_i(t)$ монотонно убывающая: она равна единице до момента отказа, а после этого момента все время остается равной нулю*.

Состояние системы полностью и однозначно определяется совокупностью состояний ее элементов. Назовем эту функциональную зависимость структурной функцией и обозначим через

$$\varphi(x) = \begin{cases} 1, & \text{если система работоспособна,} \\ 0, & \text{если система неработоспособна;} \end{cases}$$

здесь $x = (x_1, x_2, \dots, x_n)$ есть вектор состояний элементов системы, состоящей из n элементов. Для случайной булевой переменной, т. е. бернуллиевой случайной величины, можно записать

$$E\{x_i(t)\} = p_i. \quad (2.1)$$

Структурная функция надежности отображает блок-схему надежности, которая представляет собой двухполюсный граф: безот-

* Там, где это не вызовет недоразумений, для краткости будем опускать аргумент t и использовать обычную булеву переменную x_i .

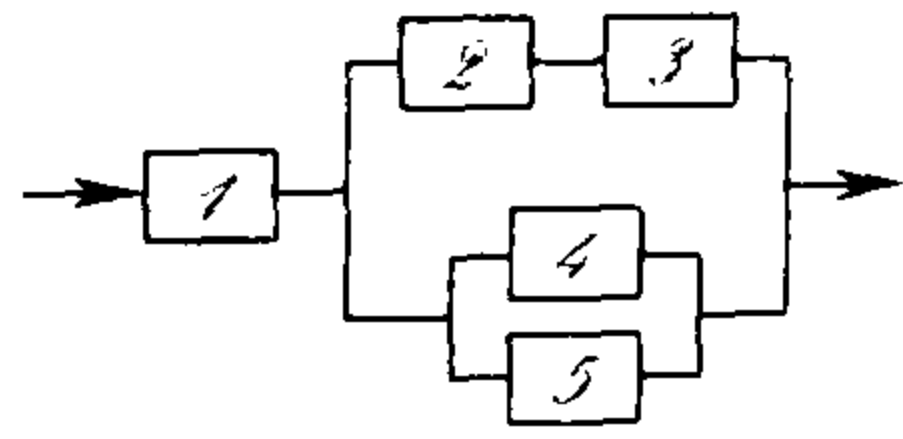


Рис. 2.1

казной работе системы соответствует *связность* этого двухполюсного графа, т. е. наличие связи «входной» вершины с «выходной». Пример блок-схемы надежности представлен на рис. 2.1.

Говорят, что система обладает свойством *монотонности* (или имеет

монотонную структуру), если

- ✓ структурная функция Φ не убывает по всем аргументам,
- ✓ $\Phi(0) = 0$ и $\Phi(1) = 1$.

Математическое ожидание структурной функции есть *вероятность связности* двухполюсного графа, отображающего систему, т. е. она эквивалентна вероятности безотказной работы системы:

$$E\{\Phi(x)\} = P_{\text{сист}}. \quad (2.2)$$

Эту же функцию называют также *функцией связности* и обозначают через

$$E\{\Phi(x)\} = h(p), \quad (2.3)$$

где $p = (p_1, p_2, \dots, p_n)$ и все p_i определены выражением (2.1).

Отметим, что функция связности $h(p)$ является монотонной функцией своих аргументов. Это имеет естественное объяснение: повышение надежности любого из элементов системы не понижает надежности системы в целом, а понижение надежности элемента не приводит к ее повышению.

§ 2.2. ПОСЛЕДОВАТЕЛЬНАЯ СИСТЕМА

Последовательной называется такая система, отказ любого элемента которой приводит к отказу системы в целом.

Последовательная система, состоящая из n элементов, описывается при помощи блок-схемы надежности, представленной на рис. 2.2.

Работоспособность любой системы в каждом ее состоянии характеризуется наличием хотя бы одного *пути*, т. е. цепочки элементов, соединяющих «вход» и «выход» системы, отображенной в виде блок-схемы надежности. Для последовательной системы такой путь является единственным.

Булева функция, удовлетворяющая этому требованию, имеет вид

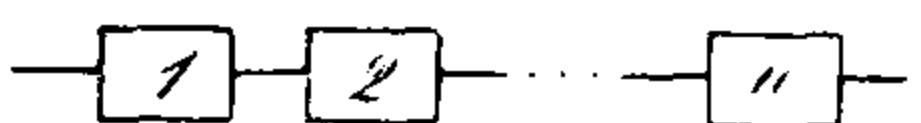


Рис. 2.2

$$\Phi(x) = \bigcap_{i=1}^n x_i. \quad (2.4)$$

Функцию (2.4) можно также записать в эквивалентной форме:

$$\Phi(x) = \max_{1 \leq i \leq n} x_i. \quad (2.5)$$

Далее мы будем предполагать элементы системы взаимно независимыми. В этом случае:

$$P_{\text{сист}} = E\{\Phi(x)\} = \prod_{i=1}^n E\{x_i\} = \prod_{i=1}^n p_i. \quad (2.6)$$

Формула (2.6) демонстрирует следующие простые свойства последовательной системы:

- ✓ с увеличением числа элементов системы ее надежность при прочих равных условиях понижается;
- ✓ надежность системы в целом всегда хуже, чем надежность любого из ее элементов (в том числе и наихудшего).

Если рассматривать поведение невосстанавливаемой системы во времени, то

$$P_{\text{сист}}(t) = \prod_{i=1}^n p_i(t). \quad (2.7)$$

Если все элементы имеют экспоненциальные распределения времени работы до отказа, то

$$P_{\text{сист}}(t) = \prod_{i=1}^n e^{-\lambda_i t} = \exp\left(-\sum_{i=1}^n \lambda_i t\right) = e^{-\Lambda t}, \quad (2.8)$$

где $\Lambda = \sum_{i=1}^n \lambda_i$.

Если элементы системы высоконадежны, т. е. $q_i(t) \ll 1/n$, где $q_i = 1 - p_i$ для всех i , то

$$P_{\text{сист}}(t) = \prod_{i=1}^n [1 - q_i(t)] \approx 1 - \sum_{i=1}^n q_i(t). \quad (2.9)$$

Среднее время работы до отказа определяется как интеграл от функции вероятности безотказной работы (ВБР). Определить значение этого интеграла в общем случае удастся лишь с помощью численных методов. Редкое исключение представляет собой экспоненциальное распределение времени работы до отказа для элементов. В этом случае из (2.8) следует:

$$T_{\text{сист}} = \int_0^{\infty} \exp\left(-t \sum_{i=1}^n \lambda_i\right) dt = \frac{1}{\sum_{i=1}^n \lambda_i} = \frac{1}{\sum_{i=1}^n \frac{1}{T_i}}, \quad (2.10)$$

где $T_i = 1/\lambda_i$ есть среднее время работы до отказа i -го элемента. Заметим, что именно случай экспоненциального распределения времени до отказа представляет наибольший практический интерес.

§ 2.3. ПРАВИЛО ДЕ МОРГАНА

Структуры систем без восстановления описываются при помощи булевых функций. Одним из важных правил работы с булевыми функциями является известное в математической логике правило де Моргана. Прежде чем перейти к нему, напомним закон двойного отрицания, который гласит, что

$$\bar{\bar{x}} = x. \quad (2.11)$$

Это правило становится очевидным, если взглянуть на цепочку диаграмм Венна (рис. 2.3).



Рис. 2.3

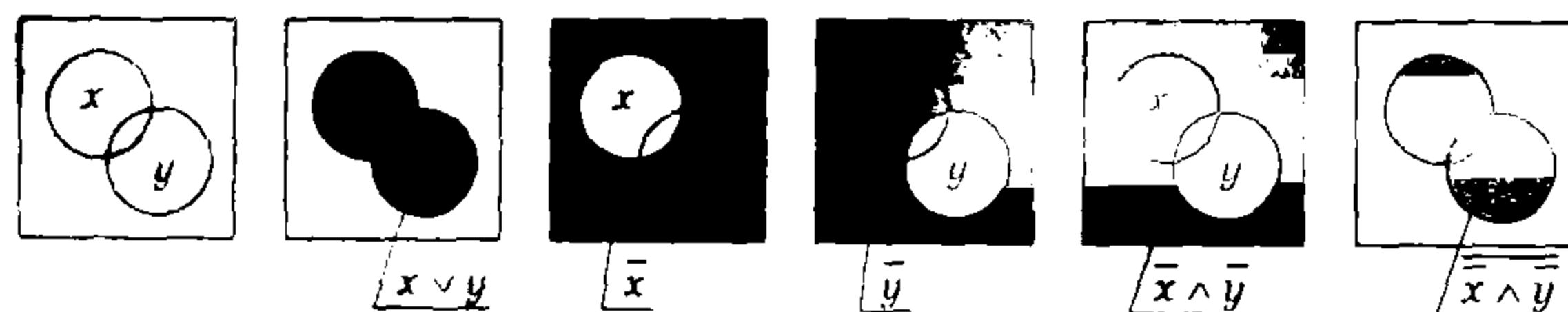


Рис. 2.4

Правило де Моргана гласит:

$$x \vee y = \overline{\bar{x} \wedge \bar{y}}. \quad (2.12)$$

Тождество (2.12) легко доказать с использованием диаграмм Венна (рис. 2.4).

Используя закон двойного отрицания, можно получить эквивалентное выражение:

$$x \wedge y = \overline{\bar{x} \vee \bar{y}}. \quad (2.13)$$

Действительно,

$$\begin{aligned} x \wedge y &= \overline{\bar{x} \vee \bar{y}} = \bar{x} \vee \bar{y} = \overline{\bar{\bar{x}} \wedge \bar{\bar{y}}} = \\ &= \bar{x} \wedge \bar{y} = \overline{\bar{x} \vee \bar{y}} = \overline{\bar{x} \vee \bar{y}}. \end{aligned} \quad (2.14)$$

Очевидно, этот же результат легко получается и с использованием диаграмм Венна.

Нетрудно аналогичным образом получить еще две эквивалентные формы:

$$\bar{x} \vee \bar{y} = \overline{x \wedge y} \quad (2.15)$$

и

$$\bar{x} \wedge \bar{y} = \overline{x \vee y}. \quad (2.16)$$

Далее, легко обобщить это правило во всех его четырех формах на любое число переменных:

$$x_1 \vee x_2 \vee \dots \vee x_n = \overline{\bar{x}_1 \wedge \bar{x}_2 \wedge \dots \wedge \bar{x}_n}, \quad (2.17)$$

$$x_1 \wedge x_2 \wedge \dots \wedge x_n = \overline{\bar{x}_1 \vee \bar{x}_2 \vee \dots \vee \bar{x}_n}, \quad (2.18)$$

$$\bar{x}_1 \vee \bar{x}_2 \vee \dots \vee \bar{x}_n = \overline{x_1 \wedge x_2 \wedge \dots \wedge x_n}, \quad (2.19)$$

$$\bar{x}_1 \wedge \bar{x}_2 \wedge \dots \wedge \bar{x}_n = \overline{x_1 \vee x_2 \vee \dots \vee x_n}. \quad (2.20)$$

Покажем принцип получения этих формул на примере формулы (2.17) для трех переменных, используя правило де Моргана и закон двойного отрицания:

$$\begin{aligned} x \vee y \vee z &= (x \vee y) \vee z = \overline{\bar{x} \wedge \bar{y}} \vee z = \\ &= \overline{\bar{x} \wedge \bar{y} \wedge \bar{z}} = \overline{\bar{x} \wedge \bar{y} \wedge \bar{z}}. \end{aligned} \quad (2.21)$$

§ 2.4. ПАРАЛЛЕЛЬНАЯ СИСТЕМА — ГОРЯЧИЙ РЕЗЕРВ

Параллельной называется система, которая сохраняет свою работоспособность до тех пор, пока работоспособен хотя бы один ее элемент. Предполагается, что все элементы находятся во включенном состоянии с самого начала.

В инженерной практике такой резерв называют *нагруженным* или *горячим*. Параллельная система, состоящая из n элементов, описывается при помощи следующей блок-схемы надежности, представленной на рис. 2.5.

В каждый момент времени функции системы выполняются основным элементом, а в случае его отказа происходит мгновенное переключение на любой из исправных к данному моменту резервных элементов.

Булева функция, удовлетворяющая этому требованию, имеет вид

$$\varphi(x) = \bigcup_{i=1}^n x_i = x_1 \cup x_2 \cup \dots \cup x_n, \quad (2.22)$$

или

$$\varphi(x) = \max_{1 \leq i \leq n} x_i. \quad (2.23)$$

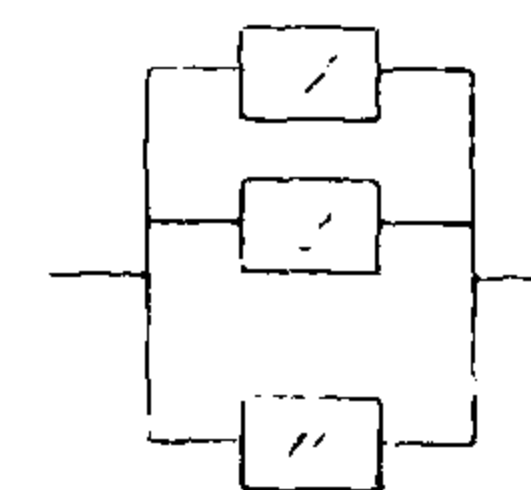


Рис. 2.5

Воспользовавшись правилом де Моргана, можно записать удобное выражение для структурной функции параллельной системы:

$$\varphi(x) = \bigcup_{i=1}^n x_i = \overline{\bigcap_{i=1}^n \bar{x}_i}, \quad (2.24)$$

из которого следует выражение для ВБР в виде

$$P_{\text{сист}} = E\{\varphi(x)\} = h(p) = 1 - \prod_{i=1}^n q_i. \quad (2.25)$$

Формула (2.25) имеет простое физическое объяснение: система отказывает при отказе всех ее элементов, а ВБР находится как вероятность дополнительного события.

Отметим следующие свойства параллельной системы:

- ✓ с увеличением числа элементов системы ее надежность при прочих равных условиях повышается;
- ✓ надежность системы в целом всегда выше, чем надежность любого из ее элементов (в том числе и наилучшего).

Обычно параллельные системы состоят из идентичных элементов. Это и естественно, так как элементы, замещающие друг друга, должны быть одинаковыми. В этом случае

$$P_{\text{сист}} = 1 - q^n. \quad (2.26)$$

Если элементы системы имеют идентичное экспоненциальное распределение времени до отказа (см. Приложение 2) и являются высоконадежными, т. е. $q(t) \ll 1$, то можно записать приближенную формулу:

$$P_{\text{сист}}(t) = 1 - (1 - e^{-\lambda t})^n \approx 1 - \lambda^n t^n \approx \exp[-(\lambda t)^n]. \quad (2.27)$$

Иначе говоря, для малых t распределение момента появления отказа параллельной системы хорошо приближается распределением Вейбулла — Гнеденко с параметрами $\alpha = \lambda$ и $\beta = n$ (см. П 2.12).

Среднее время безотказной работы параллельной системы в общем случае можно найти лишь путем численного интегрирования.

В случае экспоненциального распределения и идентичных элементов интегрирование функции вероятности безотказной работы не вызывает затруднений. Действительно, используя (2.26), можно записать:

$$T_{\text{сист}} = \int_0^{\infty} [1 - (1 - \exp(-\lambda t))^n] dt. \quad (2.28)$$

Введя новую переменную $y = 1 - \exp(-\lambda t)$, можно записать

$$T_{\text{сист}} = \frac{1}{\lambda} \int_0^1 \frac{1 - y^n}{1 - y} dy = \frac{1}{\lambda} \int_0^1 (1 + y + y^2 + \dots + y^{n-1}) dy \quad (2.29)$$

или окончательно

$$T_{\text{сист}} = \frac{1}{\lambda} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right) = \frac{1}{\lambda} \sum_{k=1}^n \frac{1}{k}. \quad (2.30)$$

Следует отметить, что среднее время безотказной работы параллельной системы с нагруженным резервом растет довольно медленно с ростом числа резервных элементов. В табл. 2.1 приведен расчет для экспоненциального распределения, которое, кстати, дает еще относительно большой рост по сравнению с большинством стареющих распределений.

В то же время мы видим, что такой вид резервирования очень эффективен для повышения ВБР.

§ 2.5. СИСТЕМА ТИПА « k ИЗ n »

Система типа « k из n » состоит из n элементов, и для ее работоспособности нужно, чтобы не менее k из них были бы работоспособны. Структурная функция такой системы может быть записана в виде

$$\varphi(x) = \begin{cases} 1, & \text{если } \sum_{i=1}^n x_i \geq k, \\ 0, & \text{если } \sum_{i=1}^n x_i < k. \end{cases}$$

Наиболее часто встречается на практике система типа «2 из 3». Для такой системы структурная функция записывается как

$$\begin{aligned} \varphi(x) = & (x_1 \cap x_2 \cap x_3) \cup (\bar{x}_1 \cap x_2 \cap x_3) \cup \\ & \cup (x_1 \cap \bar{x}_2 \cap x_3) \cup (x_1 \cap x_2 \cap \bar{x}_3). \end{aligned} \quad (2.31)$$

Здесь члены логической суммы записаны в скобках лишь для удобства прочтения.

Из (2.31) следует, что

$$P_{\text{сист}} = E\{\varphi(x)\} = p^3 + 3p^2q. \quad (2.32)$$

Таблица 2.1

Число элементов в параллельной системе	Относительный рост среднего времени работы системы
1	1
10	2.88
100	5.18
1000	6.48

В общем случае система « k из n » соответствует биномиальной схеме испытаний (см. Приложение 2), т. е. можно записать ее ВБР в виде

$$P_{\text{сист}}(t) = \sum_{j=k}^n \binom{n}{j} [p(t)]^j [q(t)]^{n-j} = 1 - \sum_{j=0}^{k-1} \binom{n}{j} [p(t)]^j [q(t)]^{n-j}. \quad (2.33)$$

Для высоконадежных систем, когда $q \ll 1/n$, имеет место простое приближенное выражение

$$P_{\text{сист}}(t) \approx 1 - \binom{n}{k-1} [q(t)]^{n-k+1}. \quad (2.34)$$

Если элементы имеют экспоненциальное распределение времени работы до отказа, то среднее время безотказной работы системы типа « k из n » находится по аналогии с (2.30) в виде

$$T_{\text{сист}} = \frac{1}{\lambda} \sum_{i=k}^n \frac{1}{i}. \quad (2.35)$$

В общем случае приходится находить $T_{\text{сист}}$, интегрируя (2.33) численными методами.

Заметим, что когда $k = n$, система типа « k из n » превращается в обычную последовательную систему

$$\varphi(x) = \prod_{i=1}^n x_i = \min_{1 \leq i \leq n} x_i, \quad (2.36)$$

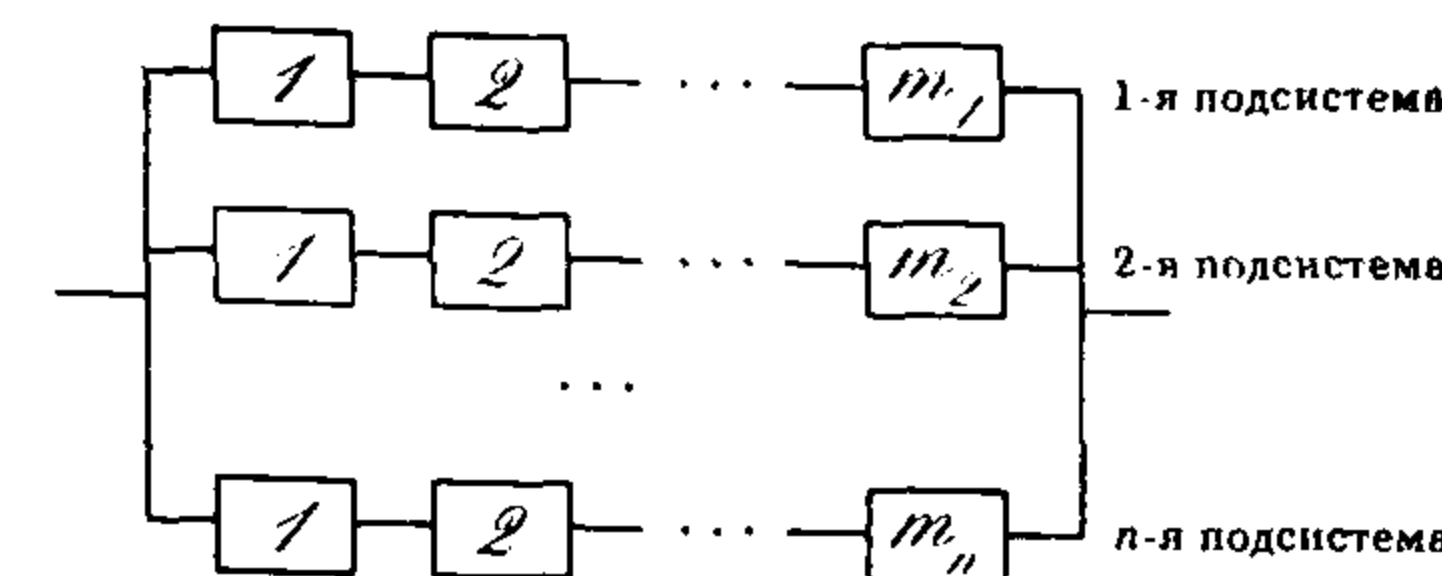
а когда $k = 1$ — в параллельную систему:

$$\varphi(x) = 1 - \prod_{i=1}^n (1 - x_i) = \max_{1 \leq i \leq n} x_i. \quad (2.37)$$

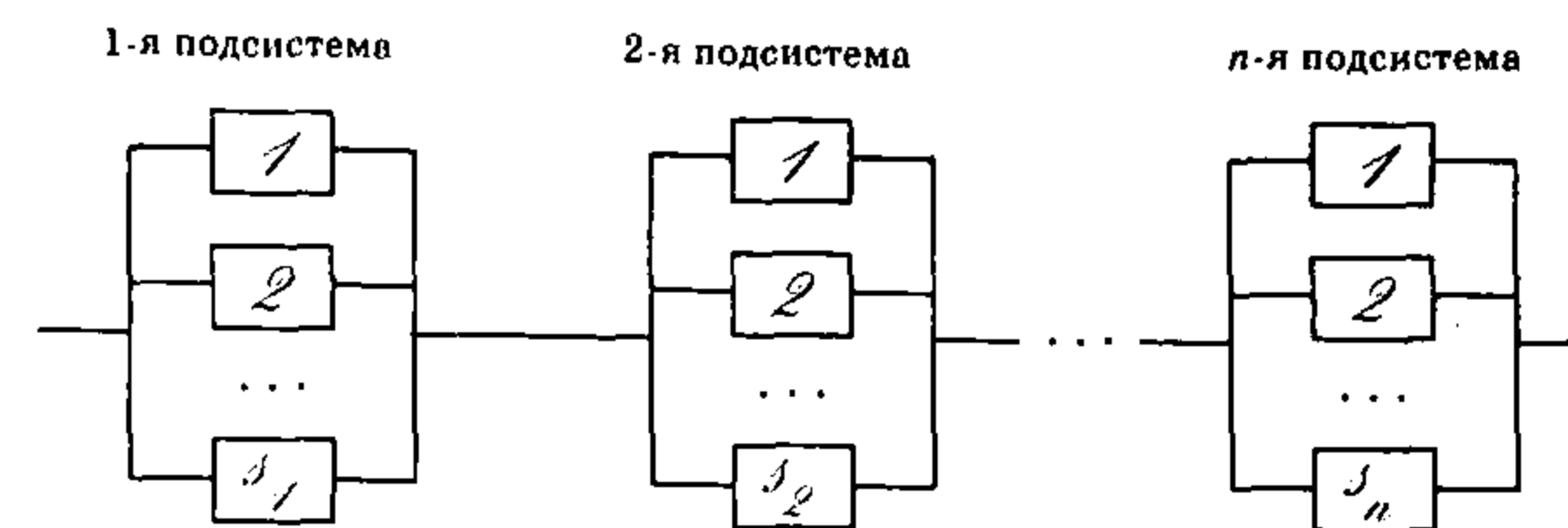
§ 2.6. СИСТЕМЫ СО СМЕШАННОЙ СТРУКТУРОЙ

2.6.1. Параллельно-последовательные и последовательно-параллельные структуры

Системы с чисто последовательной или чисто параллельной структурой встречаются на практике довольно редко. Чаще приходится иметь дело со *смешанными* структурами: с параллельно-последовательными системами (параллельное соединение последовательных подсистем) и с последовательно-параллельными системами (последовательное соединение параллельных подсистем), представленными на рис. 2.6, а также их комбинациями.



а



б

Рис. 2.6

Для параллельно-последовательной структуры, представленной на рис. 2.6, а, ВБР записывается как

$$P_{\text{сист}} = 1 - \prod_{j=1}^n \left(1 - \prod_{i=1}^{m_j} p_{ij} \right), \quad (2.38)$$

где m_j — число элементов в j -й группе последовательно соединенных элементов, а для последовательно-параллельной структуры (рис. 2.6, б) имеем

$$P_{\text{сист}} = \prod_{j=1}^n \left(1 - \prod_{i=1}^{s_j} q_{ij} \right), \quad (2.39)$$

здесь s_j — число элементов в j -й группе параллельно соединенных элементов.

2.6.2. Общий случай приводимых структур

В общем случае системы имеют гораздо более сложные смешанные структуры. Однако они почти всегда редуцируются до более простых систем, т. е. являются *приводимыми* структурами. Пояснить это удобнее всего с помощью конкретного примера. Рассмотрим смешанную структуру (рис. 2.7) и проведем с ней процедуру редукции блок-схемы надежности.

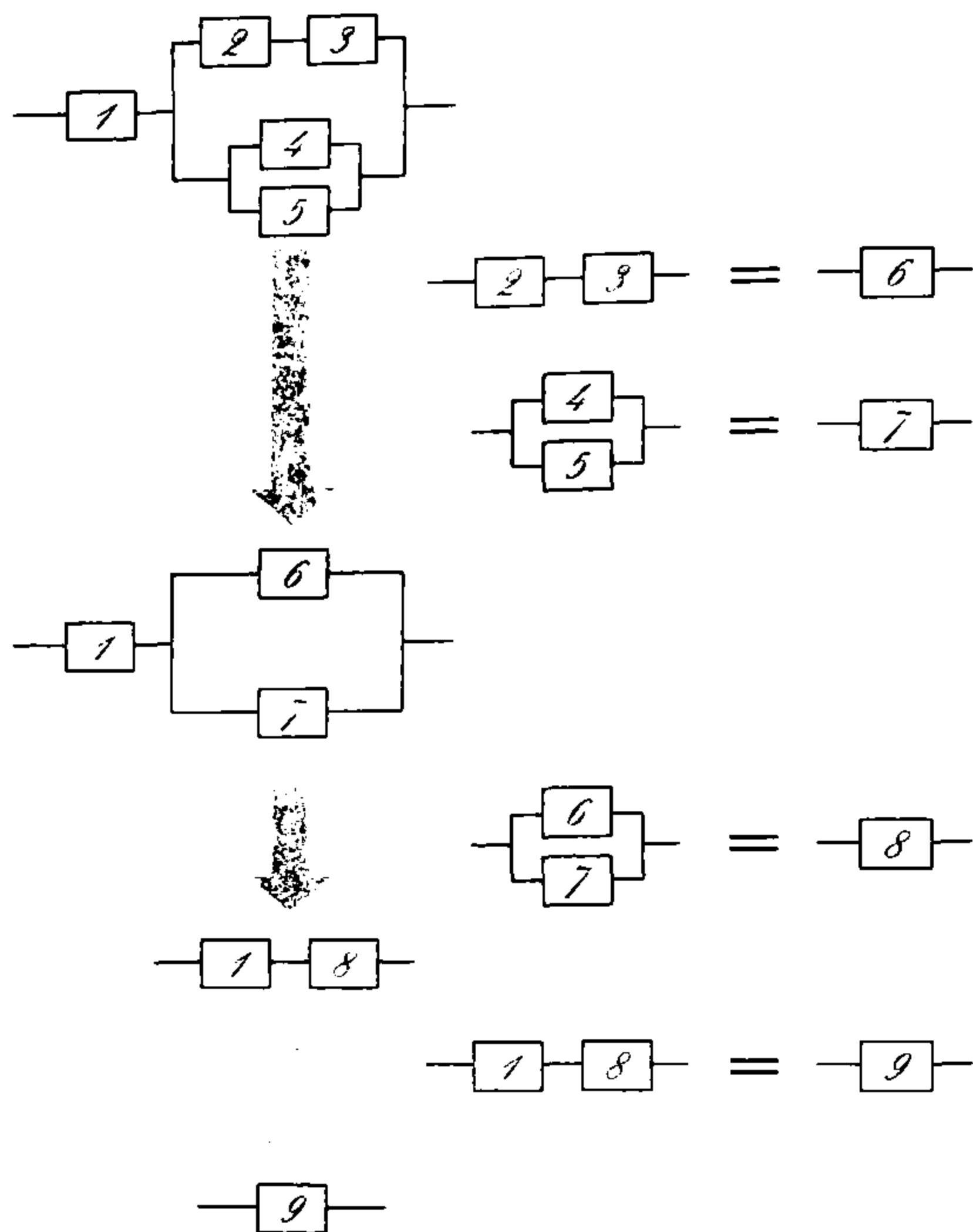


Рис. 2.7

Сначала выделим две подсистемы: последовательную из элементов 2 и 3 и параллельную из элементов 4 и 5, преобразовав эти фрагменты в соответствующие эквивалентные элементы 6 и 7. Затем элементы 6 и 7 преобразовываются в эквивалентный элемент 8. Наконец, вся система сводится к последнему эквивалентному элементу 9.

Построение выражения для вероятности безотказной работы системы начнем с конца. Ясно, что $P_{\text{сист}} = p_9$. Далее, $p_9 = p_1 p_8$. Подставляя $p_8 = 1 - q_6 q_7$, получаем $p_9 = p_1 (1 - q_6 q_7)$. Подставляя $p_6 = p_2 p_3$ и $p_7 = 1 - q_4 q_5$, получаем $p_9 = p_1 (1 - (1 - p_2 p_3) q_4 q_5)$, где искомое $P_{\text{сист}}$ выражено через вероятность безотказной работы «атомарных» элементов.

Кроме приводимых структур, имеются также неприводимые структуры, которые будут рассмотрены в главе 3.

§ 2.7. НЕНАГРУЖЕННОЕ РЕЗЕРВИРОВАНИЕ — ХОЛОДНЫЙ РЕЗЕРВ

2.7.1 Параллельная система

Уже отмечалось, что *горячее* резервирование является достаточно неэффективным с точки зрения увеличения среднего времени работы до отказа. Более эффективным является ненагруженное, или *холодное*, резервирование, когда на рабочей позиции находится единственный элемент, но в случае отказа он мгновенно заменяется на новый исправный элемент. Конечно, нужно понимать, что подобная схема мгновенной замены является идеализированной: практически подключение нового элемента требует определенного *разогрева*: в силовых электрических схемах это связано с переходными процессами в источниках питания, в механических объектах — с инерционностью подвижных деталей, а в цифровых — с необходимостью загрузки программ и повторного ввода входных данных.

Для случая *холодного* резерва условная блок-схема надежности представлена на рис. 2.8.

Рассмотрим систему, состоящую из n элементов, из которых основной находится в рабочем режиме, а остальные $(n - 1)$ — в *холодном*. Естественно предполагать, что при такой схеме резервирования все элементы системы идентичны. Также предполагается, что, находясь в резерве, элементы не теряют своих качеств и не могут отказаться. Случайное время работы системы $X_{\text{сист}}$ можно записать через соответствующие наработки отдельных элементов:

$$X_{\text{сист}} = X_1 + X_2 + \dots + X_n. \quad (2.40)$$

Для среднего времени работы системы до отказа имеем

$$T_{\text{сист}} = nT. \quad (2.41)$$

Заметим, что выражение (2.41) справедливо для любых распределений времени до отказа элементов. Распределение времени работы до отказа у такой системы определяется как распределение суммы случайных величин, т. е. выражается через n -кратную свертку:

$$P_{\text{сист}}^{(n)}(t) = 1 - \int_0^t F_{\text{сист}}^{*(n-1)}(t-x) dF(x) = 1 - F^{*n}(t), \quad (2.42)$$

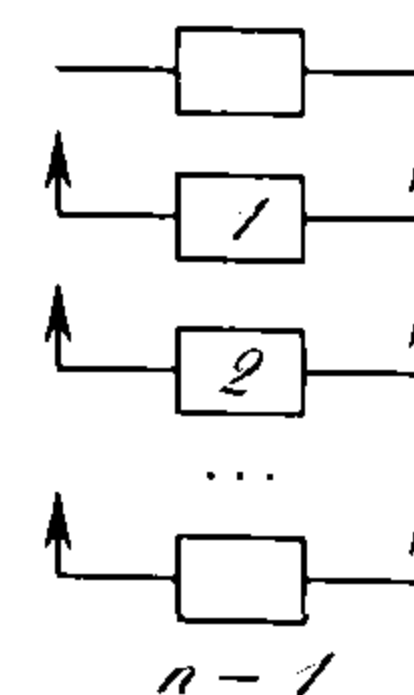


Рис. 2.8

где $F^{*n}(t)$ обозначает n -кратную свертку. По определению n -кратная свертка рекуррентно определяется как $F^{*n}(t) = F^{*(n-1)} * F(t)$ и $F^{*2}(t) = \int_0^t f(t-x) dF(x)$.

Формула (2.42) дает простое выражение в случае экспоненциальности распределения времени работы до отказа элемента. В этом случае время работы системы представляет собой сумму n экспоненциальных случайных величин с одним и тем же параметром λ .

Что означает, что рассматриваемая система не откажет за время t ? Это означает, что с вероятностью $e^{-\lambda t}$ не произойдет ни одного отказа, с вероятностью $\lambda t e^{-\lambda t}$ произойдет ровно один отказ, с вероятностью $\frac{(\lambda t)^2}{2} e^{-\lambda t}$ произойдет ровно два отказа, и так далее до $(n-1)$ отказов. Иначе говоря, ВБР такой системы записывается в виде

$$P_{\text{сист}}^{(n)}(t) = \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = 1 - \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (2.43)$$

Формула (2.43) задает распределение Пуассона (см. Приложение 2).

Из (2.43) можно получить удобную приближенную формулу для вероятности $\tilde{P}_{\text{сист}}^{(n)}(t)$ безотказной работы в случае высоконадежной системы, т. е. при $\lambda t \ll 1/n$:

$$\tilde{P}_{\text{сист}}^{(n)}(t) \approx 1 - \frac{(\lambda t)^n}{n!}, \quad (2.44)$$

причем $\tilde{P}_{\text{сист}}^{(n)}(t) > P_{\text{сист}}^{(n)}(t)$.

2.7.2 Система типа « k из n »

Система этого типа обладает резервированием более общего типа, чем предыдущая: здесь постоянно включено k элементов, а $(n-k)$

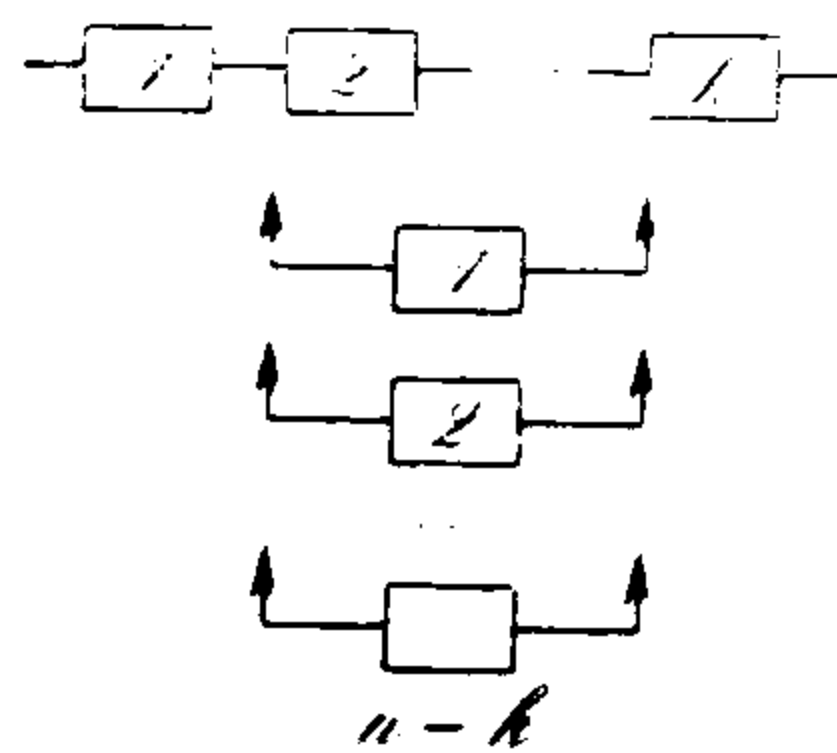


Рис 2.9

находятся в режиме ожидания. Блок-схему надежности в этом случае будем изображать в виде, представленном на рис. 2.9.

Как только возникает $(n-k+1)$ -й отказ, система в целом отказывает. Если распределение времени до отказа элемента экспоненциальное, то получить выражение для среднего времени безотказной работы и ВБР такой системы не со-

ставляет труда. Заметим, что до момента появления $(n-k+1)$ -го отказа поток отказов такой системы является пуассоновским с параметром λ_k , а следовательно:

$$T_{\text{сист}}^{(n,k)} = (n-k+1) \frac{T}{k} \quad (2.45)$$

и

$$P_{\text{сист}}^{(n,k)}(t) = \sum_{j=1}^{n-k} \frac{(\lambda k t)^j}{j!} e^{-\lambda t} = 1 - \sum_{j=n-k+1}^{\infty} \frac{(\lambda k t)^j}{j!} e^{-\lambda t}. \quad (2.46)$$

Из (2.46) следует приближенная формула для высоконадежной системы:

$$\tilde{P}_{\text{сист}}^{(n,k)}(t) \approx 1 - \frac{(\lambda k t)^{n-k+1}}{(n-k+1)!}. \quad (2.47)$$

В этом случае, как и прежде, $\tilde{P}_{\text{сист}}^{(n,k)}(t) > P_{\text{сист}}^{(n,k)}(t)$.

При произвольных распределениях времени работы до отказа элементов общие выражения абсолютно неконструктивны и обычно приходится использовать численные методы, включая метод Монте-Карло.

§ 2.8. ОБЛЕГЧЕННОЕ РЕЗЕРВИРОВАНИЕ — ТЕПЛЫЙ РЕЗЕРВ

На практике не работающие активно элементы обычно также расходуют свой ресурс, т. е. при включении в работу отличаются от новых и могут вообще отказаться, находясь еще в режиме хранения или ожидания подключения. Такой тип резервирования называют облегченным, или, на инженерном жаргоне, *теплым*. При произвольных распределениях времени работы до отказа нахождение для *теплого* резерва каких-либо показателей надежности аналитическим путем не представляется возможным, остается лишь рассчитывать на методы статистического моделирования.

Более того, модели *теплого* резервирования редко рассматриваются в инженерной практике, поскольку соответствующие исходные данные отсутствуют — не известны ни степень нагруженности, ни интенсивность отказов в режиме ожидания.

Если распределение времени безотказной работы элемента и в рабочем режиме, и в режиме хранения экспоненциальные, то задача оценки среднего времени безотказной работы имеет простое аналитическое решение:

$$T_{\text{сист}} = \sum_{j=0}^m \frac{1}{\lambda[k + (m-j)\alpha]}. \quad (2.48)$$

Глава 3

Системы с сетевой структурой

Рассматриваются общие методы анализа надежности систем с сетевой структурой, начиная с простейшей из них — мостиковой структуры. Приводятся простые иллюстративные примеры. Даются граничные оценки и приближенные формулы.

§ 3.1. МОСТИКОВАЯ СТРУКТУРА

В главе 2 рассматривались приводимые структуры. Однако на практике часто встречаются структуры, которые нельзя свести к одному эквивалентному элементу путем сворачивания последовательных и параллельных цепочек. К такого рода системам относятся практически все системы с сетевой структурой: телекоммуникационные, энергетические, транспортные. Часто рассматриваются двухполюсные системы, характеризуемые связностью входа и выхода графа, представляющего блок-схему надежности.

Мы начнем рассмотрение с мостиковой системы — простейшей системы с неприводимой структурой. Структура такой системы представлена блок-схемой надежности на рис. 3.1.

МИНИМАЛЬНЫЕ ПУТИ. Будем считать, что система работоспособна, если граф, представляющий ее на блок-схеме надежности, связан, т. е. имеется хотя бы один путь от входа до выхода двухполюсного графа.

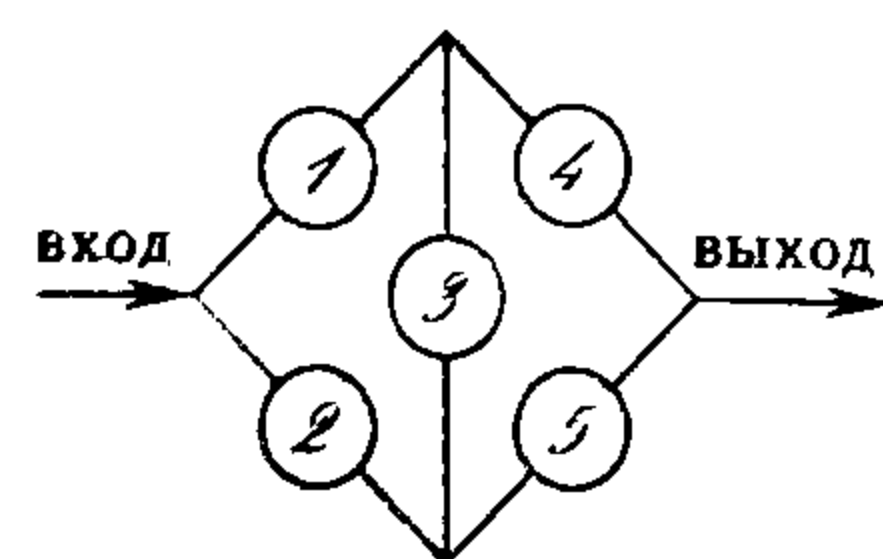


Рис. 3.1

Введем понятие минимального пути графа: это такой путь, отказ любого элемента которого приводит к потере связности данного пути, т. е. такой путь не имеет «лишних» элементов.

Заметим, что если все элементы минимального пути работоспособны, то

граф будет связан, даже если все остальные элементы отказали, но очередной отказ любого из элементов этого минимального пути приведет уже к непреклонному отказу системы. Иначе говоря, минимальный путь представляет собой последовательное соединение элементов.

В случае мостиковой системы полный набор минимальных путей представлен на рис. 3.2. Этот набор составляют подмножества элементов $\{x_1; x_4\}$, $\{x_2; x_5\}$, $\{x_1; x_3; x_5\}$ и $\{x_2; x_3; x_4\}$.

Если существует хотя бы один из указанных путей, то система работоспособна. Структурная функция мостиковой системы, выраженная через минимальные пути, записывается в виде

$$\varphi(x_1, \dots, x_5) = (x_1 \cap x_4) \cup (x_2 \cap x_5) \cup (x_1 \cap x_3 \cap x_5) \cup (x_2 \cap x_3 \cap x_4). \quad (3.1)$$

Скобки в выражении для логической суммы поставлены лишь для удобства прочтения формулы. Формулу (3.1) можно переписать, используя правило де Моргана

$$\varphi(x_1, \dots, x_5) = \overline{(\overline{x_1 \cap x_4}) \cap (\overline{x_2 \cap x_5}) \cap (\overline{x_1 \cap x_3 \cap x_5}) \cap (\overline{x_2 \cap x_3 \cap x_4})}. \quad (3.2)$$

Такая формула, содержащая только конъюнктивные формы, оказывается более удобной в дальнейших выводах.

МИНИМАЛЬНЫЕ РАЗРЕЗЫ. Очевидно, что система откажет, если произойдут отказы элементов, образующих разрез системы.

Введем понятие минимального разреза системы: это такое множество элементов системы, что при отказе всех их система отказывает, если даже все остальные элементы исправны.

В то же время восстановление работоспособности любого из элементов минимального разреза в этой же ситуации приводит к восстановлению связности двухполюсного графа. У рассматриваемой мостиковой схемы таких минимальных разрезов также

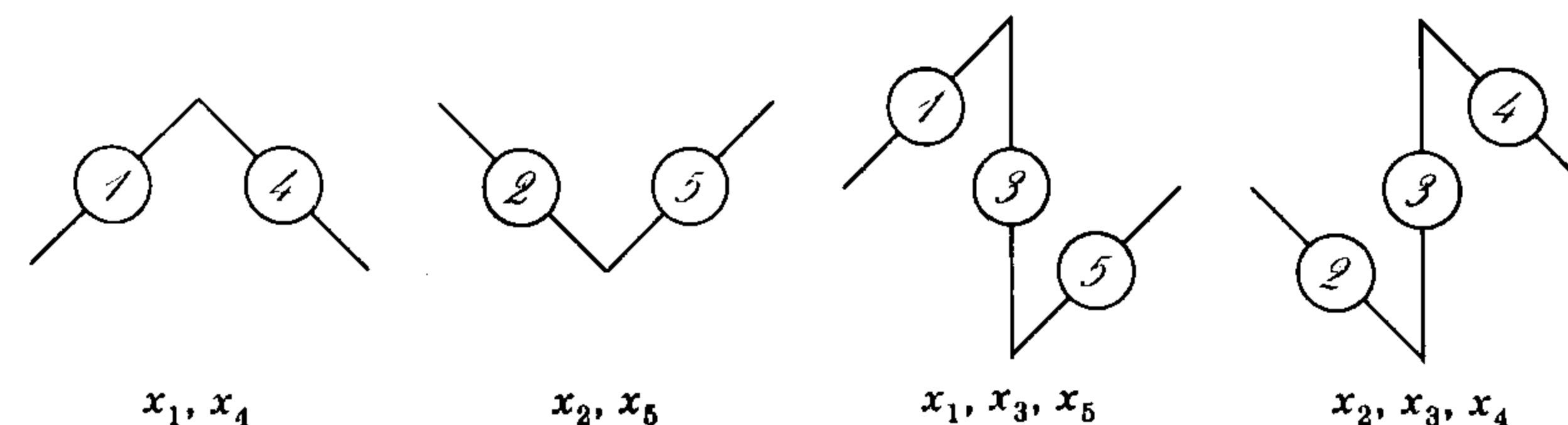


Рис. 3.2

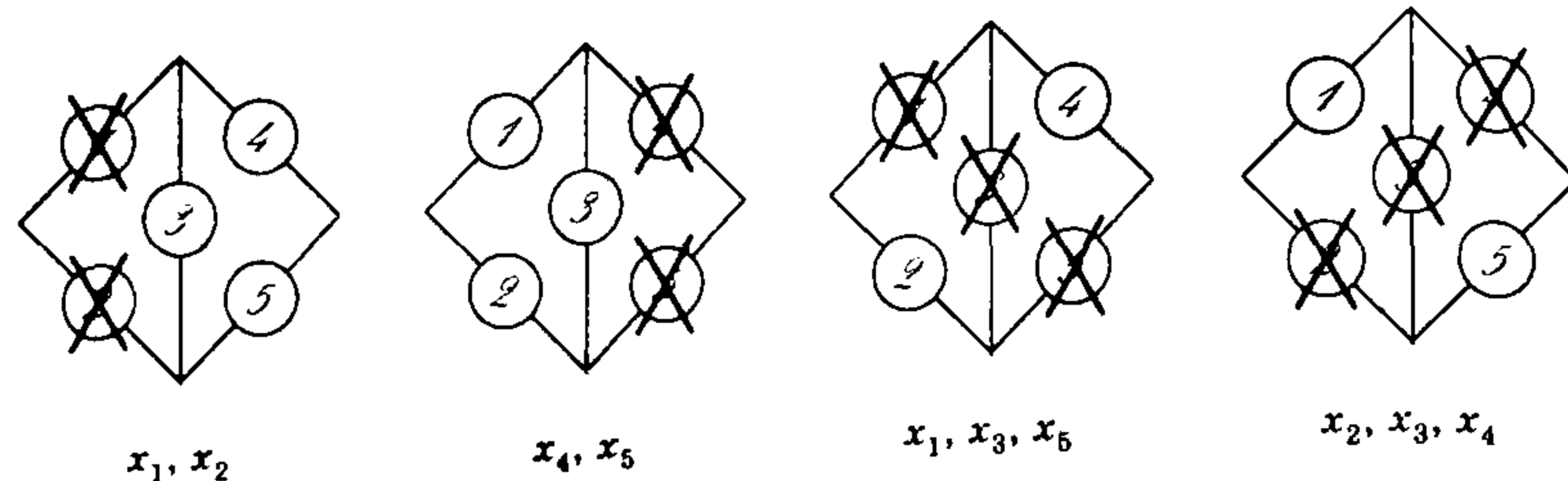


Рис. 3.3

четыре: $\{x_1; x_2\}$, $\{x_4; x_5\}$, $\{x_1; x_3; x_5\}$ и $\{x_2; x_3; x_4\}$. Эти разрезы представлены на рис. 3.3.

Булевы функции, описывающие состояние отказа перечисленных разрезов, будут иметь вид соответственно: $\bar{x}_1 \cap \bar{x}_2$, $\bar{x}_4 \cap \bar{x}_5$, $\bar{x}_1 \cap \bar{x}_3 \cap \bar{x}_5$ и $\bar{x}_2 \cap \bar{x}_3 \cap \bar{x}_4$. Отказ системы происходит, если все элементы, принадлежащие какому-либо даже только одному разрезу, откажут. Это означает, что структурная функция мостиковой системы, выраженная через минимальные разрезы, имеет вид:

$$\begin{aligned} \varphi(x_1, \dots, x_5) = \\ = \overline{(\bar{x}_1 \cap \bar{x}_2)} \cap \overline{(\bar{x}_4 \cap \bar{x}_5)} \cap \overline{(\bar{x}_1 \cap \bar{x}_3 \cap \bar{x}_5)} \cap \overline{(\bar{x}_2 \cap \bar{x}_3 \cap \bar{x}_4)}. \end{aligned} \quad (3.3)$$

Понятно, что оба представления структурной функции мостиковой системы (3.2) и (3.3) эквивалентны, что следует также из правила де Моргана.

Вероятность безотказной работы сетевой системы, или ее функция связности, есть вероятность того, что булева структурная функция равна единице, т. е.

$$P_{\text{сист}} = h(p) = h(p_1, p_2, \dots, p_n) = E\{\varphi(x_1, \dots, x_n)\}, \quad (3.4)$$

где $p_i = E\{x_i\}$ есть вероятность безотказной работы i -го элемента.

К сожалению, для сетевых структур математическое ожидание структурной функции не дает простых выражений в силу зависимости между собой путей, равно как и разрезов — из-за того, что в различные пути или разрезы могут входить одни и те же элементы.

§ 3.2. МЕТОД ВЫЧИСЛЕНИЯ ФУНКЦИИ СВЯЗНОСТИ

3.2.1. Метод прямого перебора

Наиболее простым, хотя в то же время и самым бесперспективным методом вычисления вероятности связности двухполюсного графа, является метод прямого перебора состояний. Для системы

из n элементов существует 2^n различных состояний, т. е. уже для $n = 400$ число состояний превысит гугл*. В связи с этим мы ограничимся лишь упоминанием этого метода.

3.2.2. Метод декомпозиции булевых функций

Булева функция может быть разложена относительно любого своего аргумента. Действительно, для структурной функции мостиковой системы, например, можно записать

$$\begin{aligned} \varphi(x_1, x_2, x_3, x_4, x_5) = x_3 \varphi(x_1, x_2, 1, x_4, x_5) + \\ + \bar{x}_3 \varphi(x_1, x_2, 0, x_4, x_5), \end{aligned} \quad (3.5)$$

где «1» и «0» стоят в соответствующих слагаемых на позиции 3-го элемента. В этом случае структурные функции с фиксированным значением аргумента x_3 не зависят от x_3 , поэтому вероятность связности можно записать в виде

$$\begin{aligned} P_{\text{сист}} = E\{\varphi(x_1, x_2, x_3, x_4, x_5)\} = E\{x_3\} \times E\{\varphi(x_1, x_2, 1, x_4, x_5)\} + \\ + E\{\bar{x}_3\} \times E\{\varphi(x_1, x_2, 0, x_4, x_5)\} = p_3 E\{(x_1 \cup x_2) \cap (x_4 \cup x_5)\} + \\ + q_3 E\{(x_1 \cap x_4) \cup (x_2 \cap x_5)\}, \end{aligned} \quad (3.6)$$

где $p = E\{x\}$ и $q = E\{\bar{x}\}$.

Применяя правило де Моргана, имеем

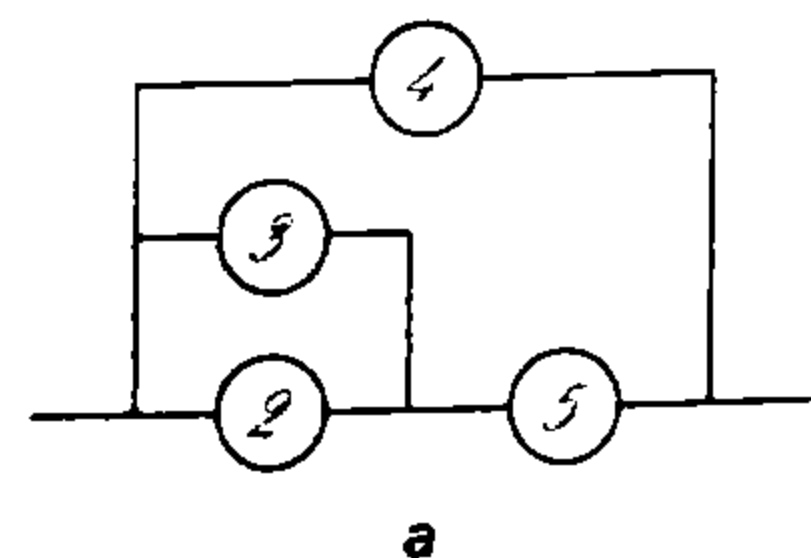
$$\begin{aligned} P_{\text{сист}} = p_3 E\{(\bar{x}_1 \cap \bar{x}_2) \cap (\bar{x}_4 \cap \bar{x}_5)\} = q_3 E\{(\bar{x}_1 \cap \bar{x}_4) \cap (\bar{x}_2 \cap \bar{x}_5)\} = \\ = p_3 (1 - q_1 q_2) (1 - q_4 q_5) + q_3 [1 - (1 - p_1 p_4) (1 - p_2 p_5)]. \end{aligned} \quad (3.7)$$

Последнюю формулу можно наглядно представить в следующей условной графической форме:

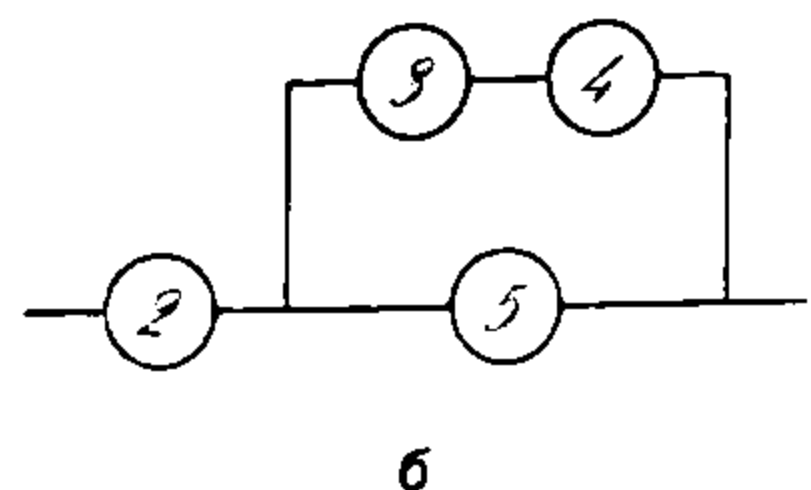
$$P\{\text{мостик}\} = p_3 P\{\text{два параллельных пути}\} + q_3 P\{\text{два последовательных элемента}\}. \quad (3.8)$$

Естественно, разложение (3.5) не является единственным: элемент, относительно которого производится разложение булевой функции, может быть и иным. В случае мостиковой схемы разложение относительно любого элемента оказывается эффективным, т. е. сводит неприводимую структуру к приводимой.

* Гугл (Google) — наибольшее физическое число, равное 10^{100} . Число элементарных частиц во Вселенной; время существования Вселенной, измеренное в наносекундах; диаметр Вселенной, измеренный в ангстремах, — все они не превышают гугла. Столь большие числа принято называть комбинаторными.



ПРИМЕР 3.1. Произведем разложение структурной функции системы с мостиковой структурой (рис. 3.1) относительно элемента 1. Две соответствующие блок-схемы надежности для случая, когда ребро 1 стянуто в точку, и для случая отсутствия ребра 1 представлены на рис. 3.4, а, б соответственно.



Расчетная формула в этом случае примет вид

$$p_1 \{1 - [1 - (1 - q_2 q_3) p_5] q_4\} + q_1 p_2 [1 - (1 - p_3 p_4) q_5].$$

Рис. 3.4

Для более сложных сетей выбор элемента, относительно которого делается разложение, не столь тривиален. Более того, может оказаться, что разложение относительно какого-то элемента вообще не упростит задачу, например если выбран элемент, уже принадлежащий к приводимой структуре.

3.2.3. Метод статистического моделирования — метод Монте-Карло

Для многих практических задач оказывается удобным применять метод Монте-Карло с использованием ЭВМ. Относительно простые задачи можно решать без программирования, используя только встроенные функции *Microsoft Excel*. Более сложные задачи требуют написания специальных программ.

МЕТОД МОДЕЛИРОВАНИЯ ОТКАЗОВ РЕБЕР. Одним из естественных методов статистического моделирования для оценки вероятности безотказной работы является имитация случайных отказов ребер и проверка связности для полученных состояний системы. В этом случае по отношению к каждому ребру сети производится следующая процедура. Выбирается случайное число v , записанное в форме десятичной дроби. Это число сравнивается с известной или заданной вероятностью безотказной работы элемента p . Если $v < p$, то элементу приписывается состояние 1, если же $v > p$, то приписывается состояние 0. После построения случайной реализации состояния всех ребер определяется, связан граф или нет. Если граф связан, то в счетчик числа успешных испытаний прибавляется единица. Иными словами, реализуется схема испыта-

ний Бернулли. После набора достаточной статистики находится частота успешных испытаний по формуле

$$\hat{P} = \frac{m}{M}, \quad (3.9)$$

где m — число успехов, а M — полное число испытаний. Естественно, что число отказов ($M - m$) должно быть достаточно большим, чтобы оценка (3.9) была достоверной.

Рассмотрим иллюстративный пример с мостиковой системой, используя для этого простейшую программу, реализованную в *Microsoft Excel*.

КРАТКОЕ ОПИСАНИЕ ПРОГРАММЫ. Вид получаемого решения приведен на рис. 3.5. Формулы, реализующие решение, представлены в табл. 3.1. Надежность элементов (т. е. ребер сети) считается в данном примере одинаковой и заносится в ячейку A5. В табл. 3.1 это значение принято равным 0.6. Решения для связности путей даны лишь для удобства и являются промежуточными. Состояние системы (ячейка A10) — это результат текущего статистического испытания.

Первое же полученное значение в ячейке A10 необходимо занести в любую пустую ячейку в *Поле решений*. При этом действии автоматически генерируются новые значения случайных чисел, и в результате появляется новое число в ячейке A10.

Процедура продолжается, пока не заполнится все поле решений. Вероятность безотказной работы (вероятность связности входа и выхода графа) считывается в ячейке F2. (Функция, реализуемая в ячейке F2: = AVERAGE (H2:R11).)

		Номера столбцов																	
		A	B	C	D	E	F	H	I	J	K	L	M	N	O	P	R		
Номера строк	1	номера элементов мостиковой схемы					РЕЗУЛЬТАТ	Поле решений											
	2	1	2	3	4	5		0.01	1	0									
	3	генератор случайных чисел						1	...										
	4	0.985405	0.777513	0.377475	0.876343	0.170366		0											
	5	0.6	индикаторы состояний элементов					1											
	6	1	1	0	1	0		0											
	7	1-й путь	2-й путь	3-й путь	4-й путь			0											
	8	1	1	1	1			1											
	9	Состояние системы						1											
	10	1						1											

Рис. 3.5

Таблица 3.1

Номера элементов мостиковой схемы				
1	2	3	4	5
Генератор случайных чисел				
=RAND()	=RAND()	=RAND()	=RAND()	=RAND()
Индикаторы состояний элементов				
=IF(A4>=\$A\$5,1,0)	=IF(B4>=\$A\$5,1,0)	=IF(C4>=\$A\$5,1,0)	=IF(D4>=\$A\$5,1,0)	=IF(E4>=\$A\$5,1,0)
1-й путь	2-й путь	3-й путь	4-й путь	
=IF(A6+D6>=1,1,0)	=IF(B6+E6>=1,1,0)	=IF(A6+C6+E6>=1,1,0)	=IF(B6+C6+D6>=1,1,0)	
Состояние системы				
=IF(A8+B8+C8+D8>=1,1,0)				

Такой метод прямого статистического моделирования оказывается весьма неэффективным в случае высоконадежных систем, поэтому ниже приводятся два других метода.

МЕТОД СТАТИСТИЧЕСКОЙ ОЦЕНКИ УСЛОВНОЙ ВЕРОЯТНОСТИ. Этот метод совмещает статистическое моделирование с аналитическим расчетом. Для простоты положим, что все ребра сети равнонадежны. Пусть также известно, что наименьший из минимальных разрезов включает k ребер, хотя число таких разрезов может быть и неизвестным.

Мы можем вычислить вероятность того, что в сети откажет ровно z ребер, по формуле

$$P_z = \binom{n}{z} p^{n-z} q^z. \quad (3.10)$$

Если бы были известны условные вероятности W_z того, что сеть окажется несвязной при условии, что в ней отказали ровно z ребер, то полная вероятность связности сети могла бы быть найдена по формуле

$$P = \sum_{z=1}^n P_z W_z. \quad (3.11)$$

Для оценки условных вероятностей W_z проведем серию статистических экспериментов, исключая из сети случайным образом ровно z

ребер, начиная с $z = k$. Статистическая оценка \hat{W}_z интересующей нас величины может быть найдена по формуле, аналогичной (3.9):

$$\hat{W}_z = \frac{m_k}{M_k}, \quad (3.12)$$

где m_k — число ситуаций, когда сеть оказывается связной после случайного извлечения из нее k ребер, а M_k — полное число испытаний. Заметим, что число испытаний для разных k может быть разным. Более того, если вероятности каких-либо состояний пренебрежимо малы, то соответствующие статистические оценки \hat{W}_z могут и не вычисляться вовсе. В результате будет получена оценка:

$$\hat{P} = \sum_{z=1}^n P_z \hat{W}_z. \quad (3.13)$$

МЕТОД СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ ВЕРОЯТНОСТИ СОСТОЯНИЙ.

В данном методе каждая реализация статистического эксперимента проводится до момента наблюдения отказа, т. е. ребра сети случайным образом исключаются одно за другим до тех пор, пока впервые не нарушится связность сети.

Пусть проведено M экспериментов и получен набор следующих пар: $(v_1, m_1), (v_2, m_2), \dots$, где v_1 — число отказавших ребер к моменту отказа сети, а m_j число таких реализаций. Все указанные величины являются случайными, хотя при этом v_1 (наименьшая из величин) не может быть меньше числа ребер в наименьшем минимальном сечении. В то же время $0 \leq m_k \leq M$.

Пусть для определенности $v_1 < v_2 < \dots$, тогда понятно, что при числе отказов ребер в сети меньшем, чем v_1 , при испытаниях не наблюдалось отказа сети в целом. Обозначим условную вероятность связности сети при условии, что в ней отказало ровно k ребер, через W_k . Тогда из того условия, что до отказа ровно v_1 ребер система ни разу не отказала, следует, что

$$\hat{W}_0 = \hat{W}_1 = \hat{W}_2 = \dots = \hat{W}_{v_1-1} = 1. \quad (3.14)$$

Затем в случае отказа $k, k > v_1$, ребер условная вероятность связности снижается до уровня:

$$\hat{W}_k = 1 - \frac{m_k}{M}. \quad (3.15)$$

Если v_s есть наибольшая из наблюдаемых величин, то, очевидно,

$$\hat{W}_{v_s} = \hat{W}_{v_s+1} = \dots = 0. \quad (3.16)$$

Итак, формула полной вероятности может быть записана в виде

$$\hat{P} = \sum_{z=0}^{v_1-1} P_z + \left(1 - \frac{v_1}{M}\right) \cdot P_{v_1} + \left(1 - \frac{v_1 + v_2}{M}\right) \cdot P_{v_2} + \\ + \left(1 - \frac{v_1 + v_2 + v_3}{M}\right) \cdot P_{v_3} + \dots + \left(1 - \frac{1}{M} \sum_{z=1}^{v_s-1} v_1\right) \cdot P_{v_s}. \quad (3.17)$$

Обычно при практических расчетах число слагаемых в сумме (3.17) бывает невелико. Формулу (3.17) можно переписать в более удобной форме:

$$\hat{P} = 1 - \frac{v_1}{M} \cdot P_{v_1} - \frac{v_1 + v_2}{M} \cdot P_{v_2} - \\ - \frac{v_1 + v_2 + v_3}{M} \cdot P_{v_3} - \dots - \left(\frac{1}{M} \sum_{z=1}^{v_s-1} v_1\right) \cdot P_{v_s-1}. \quad (3.18)$$

Последняя формула для \hat{P} , пожалуй, наиболее удобна, так как вероятности, стоящие в скобках, обычно быстро убывают. Поэтому легко увидеть, где допустимо без большой погрешности прервать вычисления.

Конечно, можно рассмотреть и общий случай, когда все ребра имеют разную надежность: это приведет лишь к более громоздким формулам, но суть рассуждений сохранится.

§ 3.3. ГРАНИЧНЫЕ ОЦЕНКИ ДЛЯ МОСТИКОВОЙ СТРУКТУРЫ

Системы из зависимых элементов

Рассматривая оценки для сетей с произвольной структурой, мы приходим к необходимости анализа последовательных и параллельных систем, элементы которых зависимы.

Рассмотрим бинарные случайные величины x_i и x_j . Из теории вероятностей известно, что если x_i и x_j коррелированы, то для вероятности их совместной реализации выполняется условие

$$P(x_i \cap x_j = 1) = P(x_i = 1) \times \\ \times P(x_j = 1) + \text{Cov}(x_i, x_j). \quad (3.19)$$

Для положительно коррелированных случайных величин имеем $\text{Cov}(x_i, x_j) > 0$, следовательно,

$$P(x_i \cap x_j = 1) > P(x_i = 1) P(x_j = 1). \quad (3.20)$$

Если случайные величины некоторой совокупности попарно положительно коррелированы (часть из них может быть независима), то они называются стохастически связанными случайными величинами. В этом случае неравенство (3.20) распространяется на произвольное число n таких связанных случайных величин. Тогда можно записать для вероятности их совместного осуществления неравенство вида:

$$P\left\{\bigcap_{j=1}^n x_j = 1\right\} = P\{x_1 \cap x_2 \cap \dots \cap x_n = 1\} \geq \prod_{j=1}^n P\{x_j = 1\} = \prod_{j=1}^n p_j, \quad (3.21)$$

где $p_i = P\{x_i = 1\}$.

Применим это правило для последовательных и параллельных систем, состоящих из указанным способом зависимых элементов. Для последовательного соединения n таких элементов из (3.21) непосредственно следует:

$$P_{\text{сист}} = P\left\{\bigcap_{j=1}^n x_j\right\} \geq \prod_{j=1}^n p_j, \quad (3.22)$$

т. е. последовательная система, состоящая из таким образом зависимых в совокупности элементов, надежнее, чем аналогичная система, состоящая из независимых элементов.

Естественно, что и для среднего времени работы до отказа будет выполняться то же условие: среднее время для последовательной системы из зависимых элементов выше, чем среднее время аналогичной системы, состоящей из независимых элементов.

Для параллельной системы из m элементов можно записать выражение для вероятности отказа:

$$Q_{\text{сист}} = 1 - P_{\text{сист}} = E\left\{\bigcap_{j=1}^m \bar{x}_j = 1\right\} \geq \prod_{j=1}^m P\{\bar{x}_j = 1\} = \prod_{j=1}^m q_j, \quad (3.23)$$

т. е. вероятность отказа параллельной системы из зависимых в упомянутом смысле элементов выше, чем вероятность отказа такой же системы из независимых элементов. Таким образом, параллельная система из независимых элементов менее надежна, чем система из зависимых элементов. Это касается как вероятности безотказной работы, так и среднего времени работы до отказа.

3.3.2. Граничные оценки Эзари — Прошана

верхняя граница. Логическая функция мостиковой системы эквивалентна логической функции параллельного соединения всех ее минимальных путей (рис. 3.6).

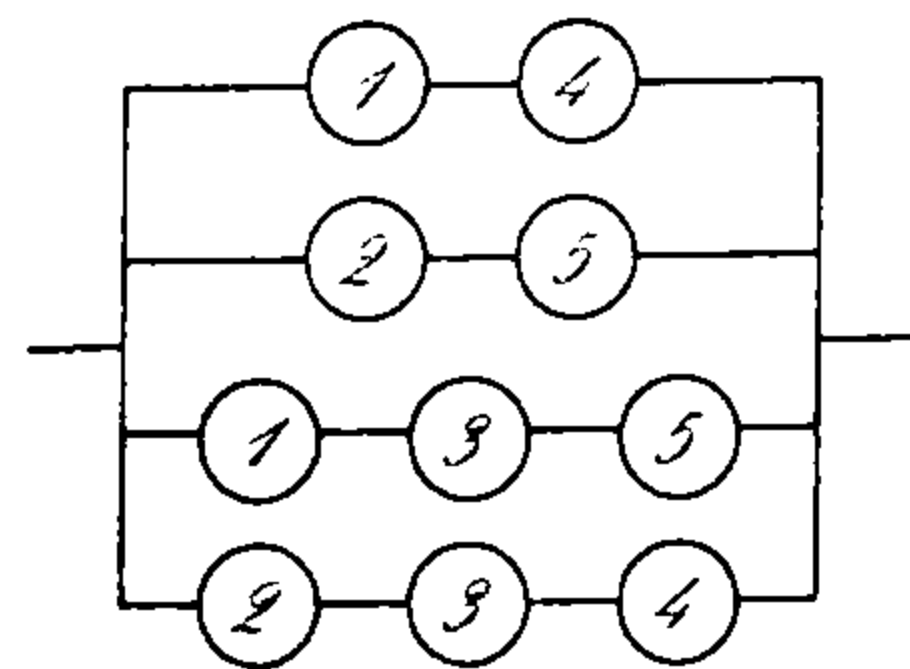


Рис. 3.6

Действительно, для того чтобы мостиковая система успешно работала, необходимо наличие хотя бы одного работоспособного пути. Схема на рис. 3.6 соответствует структурной функции (3.1). В отличие от обычной параллельно-последовательной системы параллельные цепочки на рис. 3.6 являются зависимыми, поскольку одни и те же элементы могут входить в несколько различных путей, а потому в данном случае стандартные формулы для горячего резервирования оказываются неприменимы. Пути эти зависимы таким образом, что отказ некоторого элемента k одновременно выводит из строя все пути, содержащие данный элемент.

Из толкования формулы (3.23) следует, что верхняя граница Эзари — Прошана для вероятности связности мостиковой системы равна

$$\bar{P}_{Э-П} = 1 - (1 - p_1 p_4) (1 - p_2 p_5) (1 - p_1 p_3 p_5) (1 - p_2 p_3 p_4). \quad (3.24)$$

Это же утверждение может быть проиллюстрировано условной формулой:

$$P\{\text{мостик}\} \leq P\{\text{параллельное соединение путей}\}. \quad (3.25)$$

Для высоконадежных систем, у которых $\max q_i \ll 1/n$, после элементарных преобразований (3.24) можно записать:

$$\bar{P}_{Э-П} \approx 1 - (q_1 + q_4) (q_2 + q_5) (q_1 + q_3 + q_5) (q_2 + q_3 + q_4). \quad (3.26)$$

Для идентичных элементов (3.26) примет вид:

$$\bar{P}_{Э-П} \approx 1 - 36q^4. \quad (3.27)$$

нижняя граница. Логическая функция, соответствующая выражению (3.3), может быть представлена и в виде последовательного соединения всех минимальных разрезов (рис. 3.7).

В отличие от обычной последовательно-параллельной системы последовательно включенные разрезы на рис. 3.6 являются зави-

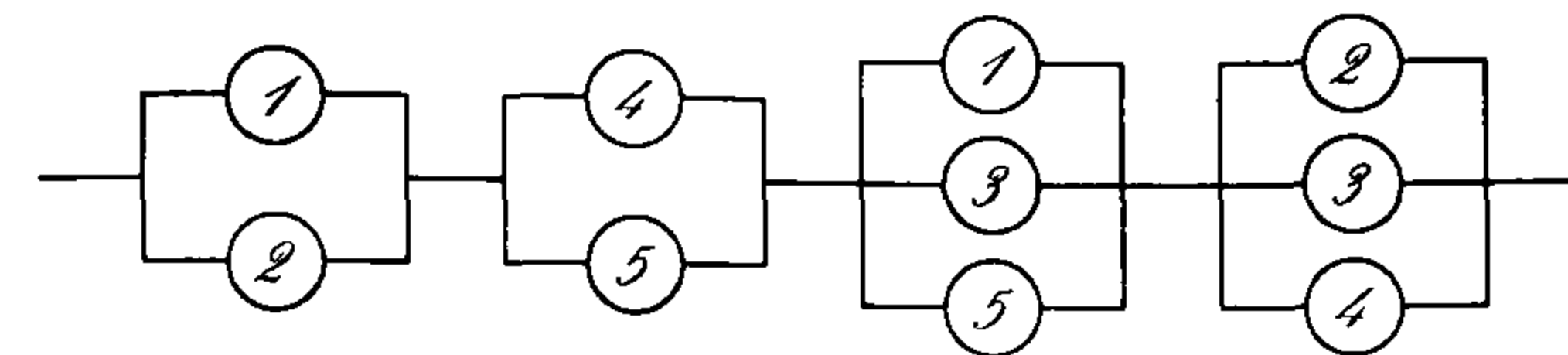


Рис. 3.7

симыми, поскольку одни и те же элементы могут входить в несколько различных разрезов. Поэтому в данном случае стандартная формула для последовательного соединения не может быть использована. Разрезы эти зависимы, причем отказ некоторого элемента k в одном из разрезов приводит к отказу одноименный элемент в другом разрезе.

В этом случае также воспользуемся результатом, приведенным в (3.22), где показано, что вероятность отказа последовательной системы из зависимых в упомянутом смысле элементов выше, чем вероятность отказа такой же системы с независимыми одноименными элементами:

$$\begin{aligned} P\{(\bar{x}_1 \cap \bar{x}_2) \cap (\bar{x}_4 \cap \bar{x}_5) \cap (\bar{x}_1 \cap \bar{x}_3 \cap \bar{x}_5) \cap (\bar{x}_2 \cap \bar{x}_3 \cap \bar{x}_4) = 1\} &\geq \\ &\geq P\{(\bar{x}_1 \cap \bar{x}_2) = 1\} \times P\{(\bar{x}_4 \cap \bar{x}_5) = 1\} \times P\{(\bar{x}_1 \cap \bar{x}_3 \cap \bar{x}_5) = 1\} \times \\ &\times P\{(\bar{x}_2 \cap \bar{x}_3 \cap \bar{x}_4) = 1\}. \end{aligned} \quad (3.28)$$

Подставляя вероятности $q_k = P\{(\bar{x}_k = 1)\}$, из (3.28) находим нижнюю границу Эзари — Прошана для $\underline{P}_{Э-П}$ мостиковой системы:

$$\underline{P}_{Э-П} \leq (1 - q_1 q_2) (1 - q_4 q_5) (1 - q_1 q_3 q_5) (1 - q_2 q_3 q_4). \quad (3.29)$$

Формула (3.29) может быть также записана в условной форме:

$$P\{\text{мостик}\} \geq P\{\text{последовательное соединение разрезов}\}. \quad (3.30)$$

Для высоконадежных систем, у которых $\max q_i \ll 1/n$, после элементарных преобразований (3.30) можно записать

$$\underline{P}_{Э-П} \approx 1 - q_1 q_2 - q_4 q_5, \quad (3.31)$$

что для идентичных элементов дает приближение

$$\underline{P}_{Э-П} \approx 1 - 2q_2. \quad (3.32)$$

Естественно, что при $\max q_i \rightarrow 0$ верхняя и нижняя границы сближаются, стремясь к единице.

Таким образом, получены двусторонние оценки Эзари — Прошана для вероятности связности мостиковой системы:

$$(1 - q_1 q_2)(1 - q_4 q_5)(1 - q_1 q_3 q_5)(1 - q_2 q_3 q_4) \leq P_{\text{сст}} \leq 1 - (1 - p_1 p_4)(1 - p_2 p_5)(1 - p_1 p_3 p_5)(1 - p_2 p_3 p_4). \quad (3.33)$$

3.3.3. Граничные оценки Литвака — Ушакова

При расчете границ Литвака — Ушакова используются только реберно-непересекающиеся минимальные пути и минимальные разрезы сетевой системы, т. е. подмножества элементов такие, в которых не встречается какой-либо элемент дважды.

нижняя граница. Сначала рассмотрим представления в виде различных наборов независимых путей. Для мостиковой системы может быть составлено три таких варианта (рис. 3.8).

Поясним процедуру построения набора минимальных путей для мостиковой системы (рис. 3.1).

(а) Выделим путь (x_1, x_4) . Тогда в графе остаются элементы x_2 , x_3 и x_5 . Элемент x_3 — висячее ребро, т. е. является несущественным. Оставшиеся элементы образуют еще один независимый путь (x_2, x_5) . Итак, образован первый набор независимых минимальных путей: (x_1, x_4) и (x_2, x_5) .

(б) Выделим путь (x_1, x_3, x_5) . После этого в сети остались только элементы, соответствующие изолированному висячему ребру, т. е. построение второго набора независимых путей закончено.

(в) Выделим путь (x_2, x_3, x_4) и повторим рассуждения, приведенные в пункте (б).

Таким образом, построены все возможные наборы независимых (реберно-непересекающихся) путей.

Прежде чем двигаться дальше, напомним, что сетевые структуры обладают естественным свойством монотонности: повышение надежности какого-либо элемента приводит к повышению

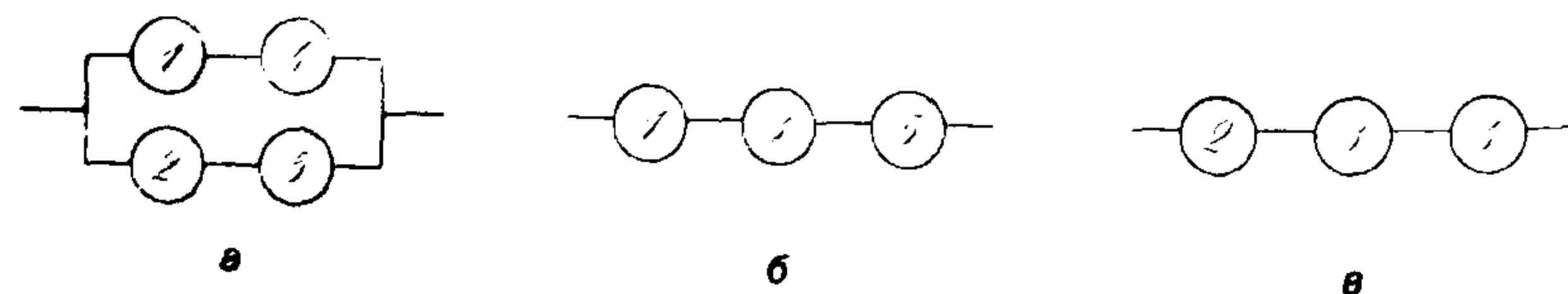


Рис. 3.8

надежности системы в целом, а понижение надежности элемента приводит, соответственно, к понижению надежности системы. Действительно, вариант (а) на рис. 3.8 соответствует ситуации, когда элемент x_3 абсолютно ненадежен (исключен); вариантам (б) и (в) соответствуют исключение элементов $\{x_2, x_4\}$ и $\{x_1, x_5\}$ соответственно. Во всех этих случаях надежность системы ухудшена. Иными словами, надежность всех вариантов, представленных на рис. 3.8, может рассматриваться как нижняя граница неизвестного истинного значения. Из нижних оценок лучшей является наибольшая из них, поэтому для мостиковой системы можно записать следующее ее значение:

$$P_{\text{л.у}} = \max\{P_{\text{л.у}}^{(a)}, P_{\text{л.у}}^{(б)}, P_{\text{л.у}}^{(в)}\} = \max\{1 - (1 - p_1 p_4)(1 - p_2 p_5), p_1 p_3 p_5, p_2 p_3 p_4\}. \quad (3.34)$$

Для высоконадежных систем, т. е. при $\max q_i \ll 1/n$, нижняя граница в данном случае будет иметь вид

$$P_{\text{л.у}} \approx \max\{1 - (q_1 + q_4)(q_2 + q_5); 1 - (q_1 + q_3 + q_5); 1 - (q_2 + q_3 + q_4)\}, \quad (3.35)$$

что обычно (при близких значениях q_i) сводит (3.35) к выражению

$$P_{\text{л.у}} \approx 1 - (q_1 + q_4)(q_2 + q_5). \quad (3.36)$$

верхняя граница. Теперь перейдем к построению верхней границы Литвака — Ушакова. Рассмотрим все подмножества независимых (реберно-непересекающихся) разрезов мостиковой структуры, представленной на рис. 3.1. Всего имеется три таких представления.

Вариант на рис. 3.9, а отвечает ситуации, когда элемент x_3 абсолютно надежен; варианты рис. 3.9, б, в — случаям, когда элементы $\{x_2, x_4\}$ и $\{x_1, x_5\}$ соответственно абсолютно надежны. Во всех этих случаях надежность системы улучшена, причем надежность каждого из вариантов, представленных на рис. 3.8, может

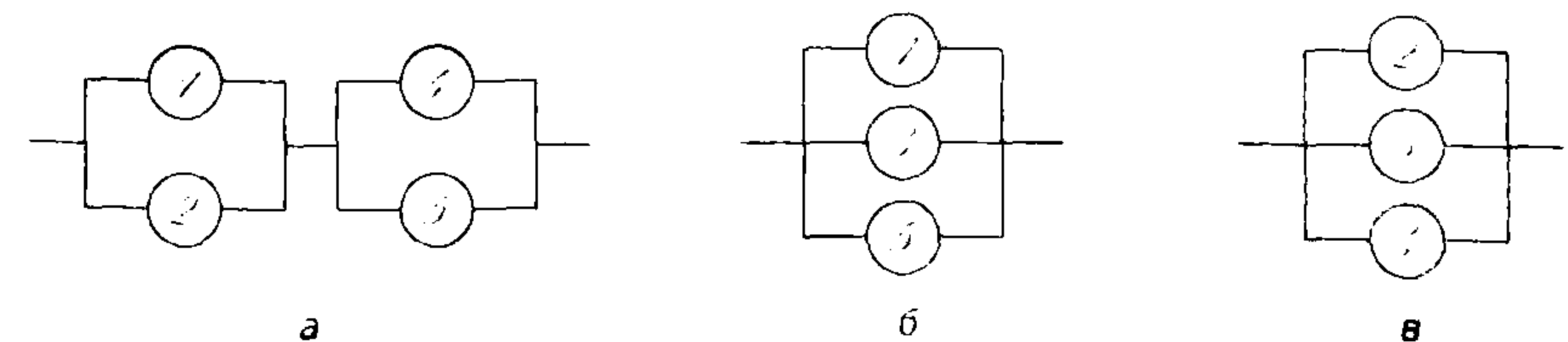


Рис. 3.9

рассматриваться как верхняя граница истинного значения. Очевидно, лучшей из верхних оценок является наименьшая. Иными словами, применительно к мостиковой системе можно записать:

$$\bar{P}_{Л-У} = \min\{\bar{P}_{Л-У}^{(a)}, \bar{P}_{Л-У}^{(b)}, \bar{P}_{Л-У}^{(c)}\} = \min\{(1 - q_1 q_2)(1 - q_4 q_5); 1 - q_1 q_3 q_5; 1 - q_2 q_3 q_4\}. \quad (3.37)$$

При $q_i \ll 1/n$ нижняя оценка будет иметь вид

$$\bar{P}_{Л-У} \approx \max\{1 - (q_1 q_2 + q_5 q_4); 1 - q_1 q_3 q_5; 1 - q_2 q_3 q_4\}, \quad (3.38)$$

что обычно (при близких значениях q_i) сводит (3.38) к выражению

$$\bar{P}_{Л-У} \approx \max\{1 - q_1 q_3 q_5; 1 - q_2 q_3 q_4\}, \quad (3.39)$$

или для идентичных элементов

$$\bar{P}_{Л-У} \approx 1 - q^3. \quad (3.40)$$

Таким образом, получены верхняя и нижняя границы Литвака — Ушакова для мостиковой системы:

$$\max\{1 - (1 - p_1 p_4)(1 - p_2 p_5), p_1 p_3 p_5, p_2 p_3 p_4\} \leq P_{\text{сист}} \leq \min\{(1 - q_1 q_2)(1 - q_4 q_5), 1 - q_1 q_3 q_5, 1 - q_2 q_3 q_4\}. \quad (3.41)$$

3.3.4. Сравнение граничных оценок Эзари — Прошана и Литвака — Ушакова

В табл. 3.2 приведены результаты вычислений верхних и нижних границ, полученные методами Эзари — Прошана и Литвака — Ушакова. Наиболее важными являются границы для значений вероятности безотказной работы элементов, близких к единице. На рис. 3.10 представлены графики зависимости верхних и нижних границ надежности мостиковой системы от надежности ее элементов. Элементы предполагаются идентичными.

На рис. 3.10 обозначены соответственно: жирными линиями 1 и 2 — верхняя и нижняя оценки Эзари — Прошана, а тонкими 3 и 4 — верхняя и нижняя оценки Литвака — Ушакова. Заметим, что для практики важнее случай относительно высоких p , т. е. оценка Эзари — Прошана оказывается в данном случае

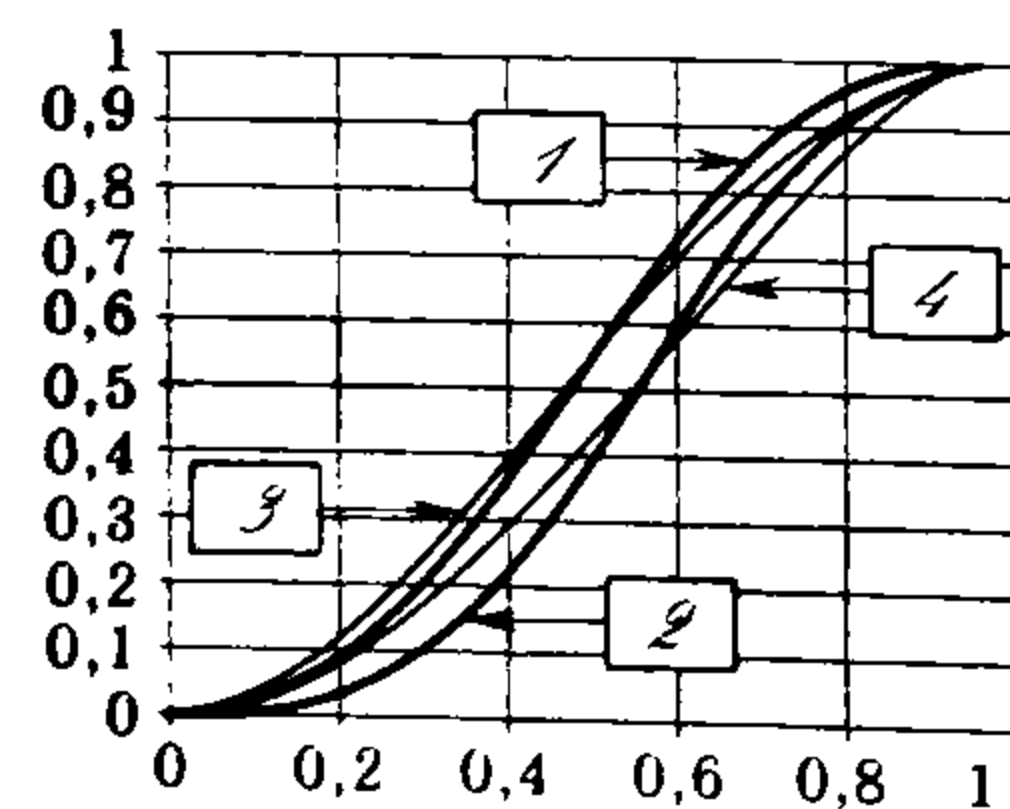


Рис. 3.10

более предпочтительной. Вычислительные аспекты, связанные с нахождением оценок по методу Эзари — Прошана, будут обсуждены ниже в § 3.4 при рассмотрении произвольных двухполусетевых систем.

§ 3.4. СЕТЕВЫЕ СИСТЕМЫ ОБЩЕГО ВИДА

Выше была рассмотрена простейшая неприводимая система — мостиковая. Однако на практике встречаются обычно гораздо более сложные системы с сетевой структурой — телекоммуника-

Таблица 3.2

P	Верхняя Л—У	Верхняя Э—П	Точное	Нижняя Л—У	Нижняя Э—П
0	0	0	0	0	0
0.02	0.0016	0.0008	0.0008	0.0008	0
0.04	0.0061	0.0033	0.0033	0.0032	0.0001
0.06	0.0135	0.0076	0.0076	0.0072	0.0004
0.08	0.0236	0.0138	0.0137	0.0128	0.0012
0.10	0.0361	0.0219	0.0215	0.0199	0.0027
0.20	0.1296	0.0931	0.0886	0.0784	0.0309
0.30	0.2601	0.216	0.1984	0.1719	0.1123
0.40	0.4096	0.3818	0.3405	0.2944	0.2518
0.50	0.5625	0.5693	0.5000	0.4375	0.4307
0.60	0.7056	0.7482	0.6595	0.5904	0.6182
0.70	0.8281	0.8877	0.8016	0.7399	0.7840
0.80	0.9216	0.9691	0.9114	0.8704	0.9069
0.90	0.9801	0.9973	0.9785	0.9639	0.9781
0.92	0.9872	0.9988	0.9863	0.9764	0.9862
0.94	0.9928	0.9996	0.9924	0.9865	0.9924
0.96	0.9968	0.9999	0.9967	0.9939	0.9967
0.98	0.9992	1.0000	0.9992	0.9984	0.9992

более предпочтительной. Вычислительные аспекты, связанные с нахождением оценок по методу Эзари — Прошана, будут обсуждены ниже в § 3.4 при рассмотрении произвольных двухполусетевых систем.

§ 3.4. СЕТЕВЫЕ СИСТЕМЫ ОБЩЕГО ВИДА

Выше была рассмотрена простейшая неприводимая система — мостиковая. Однако на практике встречаются обычно гораздо более сложные системы с сетевой структурой — телекоммуника-

ционные, энергетические, транспортные. Для этих систем могут быть применены те же методы оценки вероятности связности двухполюсных сетей, которые были продемонстрированы выше.

3.4.1. Метод декомпозиции булевых функций

В общем случае для сложных сетевых систем метод разложения булевой функции относительно ее аргументов на практике оказывается неэффективным. Можно по-прежнему записать для структурной функции выражение

$$\begin{aligned} \varphi(x_1, \dots, x_k, \dots, x_n) &= x_k \varphi(x_1, \dots, 1, \dots, x_n) + \\ &+ \bar{x}_k \varphi(x_1, \dots, 0, \dots, x_n), \end{aligned} \quad (3.42)$$

где через 1_k и 0_k обозначены единица и нуль, стоящие в соответствующих слагаемых на k -й позиции. В этом случае структурная функция с фиксированным значением аргумента не зависит от этого аргумента, поэтому

$$\begin{aligned} P_{\text{сист}} &= E\{\varphi(x_1, \dots, x_k, \dots, x_n)\} = E\{x_k\} \times E\{\varphi(x_1, \dots, 1, \dots, x_n)\} + \\ &+ E\{\bar{x}_k\} \times E\{\varphi(x_1, \dots, 0, \dots, x_n)\} = p_k E\{\varphi(x_1, \dots, 1, \dots, x_n)\} + \\ &+ q_k E\{\varphi(x_1, \dots, 0, \dots, x_n)\}. \end{aligned} \quad (3.43)$$

Однако само нахождение обеих величин $\{\varphi(x_1, \dots, 1, \dots, x_n)\}$ и $\{\varphi(x_1, \dots, 0, \dots, x_n)\}$ по-прежнему остается труднорешаемой задачей.

Естественно, разложения (3.42) и (3.43) не являются единственными: элементы, относительно которых производится разложение булевой функции, могут быть любыми иными. Кроме того, можно разложить булеву функцию и по нескольким аргументам.

Для сложных сетей, как уже отмечалось, выбор элемента, относительно которого делается разложение, нетривиален. Так, сетевая структура может включать в свой состав фрагменты, представляющие собой приводимые структуры. Если в качестве элемента, относительно которого производится разложение, выбран один из элементов приводимых структур, то это фактически не приводит ни к какому упрощению вычислений. Более того, в сложных сетевых структурах такое разложение приходится производить неоднократно, а в этом случае правильный выбор элементов весьма существен, иначе решение будет сведено к абсолютно неконструктивному методу прямого перебора.

Заметим, что задача еще более усложнится, если будут рассматриваться сети, представляемые непланарными графами.

3.4.2. Граничные оценки Эзари — Прошана

Оценки Эзари — Прошана были приведены выше для простейшего случая — мостиковой структуры. Здесь приводится описание этих оценок на общий случай.

ВЕРХНЯЯ ГРАНИЦА. Обозначим через $\alpha(X)$ структурную функцию i -го пути. Для произвольной сетевой системы с N различными минимальными путями можно записать структурную функцию в виде дизъюнкции логических функций всех минимальных путей сетевой системы:

$$\varphi(X) = \bigcup_{j=1}^N \alpha_j(X) = \overline{\bigcap_{j=1}^N \overline{\alpha_j(X)}}, \quad (3.44)$$

где последнее преобразование проведено с использованием правила де Моргана. Отсюда следует:

$$\overline{\varphi(X)} = \bigcap_{j=1}^N \overline{\alpha_j(X)}. \quad (3.45)$$

Для связанных случайных событий вероятность их совместного осуществления выше, чем произведение вероятностей каждого из событий. Это позволяет записать:

$$E\{\overline{\varphi(X)}\} \geq \prod_{j=1}^N E\{\overline{\alpha_j(X)}\}. \quad (3.46)$$

Выражение (3.46) означает, что вероятность отказа сетевой системы выше, чем вероятность отказа параллельного соединения всех независимых минимальных путей. Это утверждение эквивалентно тому, что вероятность связности сетевой системы при тех же условиях ниже, чем у системы, составленной из параллельного соединения полного набора независимых минимальных путей. Таким образом, получена верхняя оценка Эзари — Прошана для вероятности связности системы с сетевой структурой. Эту границу можно переписать в виде

$$\bar{P}_{\text{э-п}} = 1 - \prod_{j=1}^N \left(1 - P\{\alpha_j(X) = 1\}\right) = 1 - \prod_{j=1}^N \left(1 - \prod_{i \in \alpha_j} p_{ij}\right), \quad (3.47)$$

где α_j — множество индексов элементов, входящих в путь $\alpha_j(X)$.

НИЖНЯЯ ГРАНИЦА. Обозначим через $\beta(X)$ структурную функцию j -го разреза. Структурную функцию сетевой системы можно выразить как логическое произведение (конъюнкцию) структурных функций для всех минимальных разрезов системы в виде

$$\varphi(X) = \bigcap_{j=1}^M \beta_j(X), \quad (3.48)$$

где M — полное число минимальных разрезов в двухполюсной сетевой системе. Используя опять неравенство для вероятности одновременной реализации связанных случайных событий, можно записать

$$E\{\varphi(X)\} \geq \prod_{j=1}^M E\{\beta_j(X)\}, \quad (3.49)$$

т. е.

$$P_{\text{Э.П.}} = \prod_{j=1}^M \left(1 - \prod_{i \in \beta_j} q_i\right), \quad (3.50)$$

где β_j — множество индексов элементов, входящих в разрез $\beta_j(X)$.

Таким образом, получены двусторонние оценки Эзари — Прошана для вероятности связности произвольной двухполюсной сетевой системы:

$$P_{\text{Э.П.}} \leq P_{\text{сист}} \leq \bar{P}_{\text{Э.П.}} \quad (3.51)$$

В раскрытой форме эти оценки имеют вид

$$\prod_{j=1}^M \left(1 - \prod_{i \in \beta_j} q_i\right) \leq P_{\text{сист}} \leq 1 - \prod_{j=1}^N \left(1 - \prod_{i \in \alpha_j} p_{ij}\right). \quad (3.52)$$

Граничные оценки Литвака — Ушакова

Эти оценки используют только реберно-непересекающиеся минимальные пути и минимальные разрезы сетевой системы, т. е. такие подмножества элементов, в которых одни и те же элементы не встречаются дважды.

Сетевые системы, как уже отмечалось, обладают естественным свойством монотонности: повышение надежности любого элемента приводит к повышению надежности системы, а понижение надежности любого элемента приводит, соответственно, к понижению надежности системы. Идея предлагаемых оценок базируется на этом свойстве монотонных структур.

нижняя оценка Выделим в сетевой системе различные наборы взаимно независимых минимальных путей, а остальные элементы исключим из сети, понизив их надежность до нуля, понижая тем самым и надежность сетевой системы в целом.

Если в сетевой системе выделен некоторый набор независимых минимальных путей, индексы которых образуют множество A_k , то для каждого набора A_k можно записать следующее неравенство:

$$\varphi_k(X) = \bigcup_{j \in A_k} \alpha_j(X), \quad (3.53)$$

где $\alpha_j(X) = \bigcap_{i \in \alpha_j} x_i$.

Используя правило де Моргана и взяв математические ожидания от обеих частей неравенства (3.53), получаем

$$\begin{aligned} E\{\varphi_k(X)\} &= E\left\{\bigcup_{j \in A_k} \alpha_j(X)\right\} = E\left\{\overline{\bigcap_{j \in A_k} \overline{\alpha_j(X)}}\right\} = \\ &= 1 - \prod_{j \in A_k} \left(1 - \prod_{i \in \alpha_j(X)} p_{ij}\right). \end{aligned} \quad (3.54)$$

Пусть всего можно выбрать N различных подмножеств

$$A_1, A_2, \dots, A_N.$$

По построению, для любого из наборов k выполняется условие

$$\varphi(X) \geq \varphi_k(X),$$

т. е.

$$\varphi(X) \geq \max\{\varphi_1(X), \varphi_2(X), \dots, \varphi_N(X)\}. \quad (3.55)$$

Применив оператор математического ожидания к обеим частям неравенства (3.55), получаем

$$P_{\text{Л.У.}} = \max\{E\{\varphi_1(X)\}, E\{\varphi_2(X)\}, \dots, E\{\varphi_N(X)\}\}. \quad (3.56)$$

В результате можно записать:

$$\begin{aligned} P_{\text{Л.У.}} &= \max\left\{1 - \prod_{\alpha_j \in A_1} \left(1 - \prod_{i \in \alpha_j} p_{ij}\right), \right. \\ &\quad \left. 1 - \prod_{\alpha_j \in A_2} \left(1 - \prod_{i \in \alpha_j} p_{ij}\right), \dots, 1 - \prod_{\alpha_j \in A_N} \left(1 - \prod_{i \in \alpha_j} p_{ij}\right)\right\}. \end{aligned} \quad (3.57)$$

Для построения таких наборов независимых минимальных путей может быть предложен следующий эвристический алгоритм, дающий хорошие практические результаты.

АЛГОРИТМ ПОСТРОЕНИЯ НАБОРОВ НЕЗАВИСИМЫХ МИНИМАЛЬНЫХ ПУТЕЙ. Поскольку нас интересует наилучшая из нижних границ, построение начинается с кратчайших минимальных путей. Заметим при этом, что метрикой длины может быть не только число ребер в последовательной цепочке, но и максимальная надежность цепочки (т. е. *кратчайшим* путем может быть назван максимально надежный путь).

Проиллюстрируем процедуру построения независимых минимальных путей сетевой системы с помощью примера.

ПРИМЕР 3.2. Пусть мы имеем сеть, представленную на рис. 3.11, а. На рис. 3.11, б показан первый наикратчайший путь. На рис. 3.11, в построен условный наикратчайший путь в оставшемся графе при условии исключения всех ребер первого выбранного пути. На рис. 3.11, г построен следую-

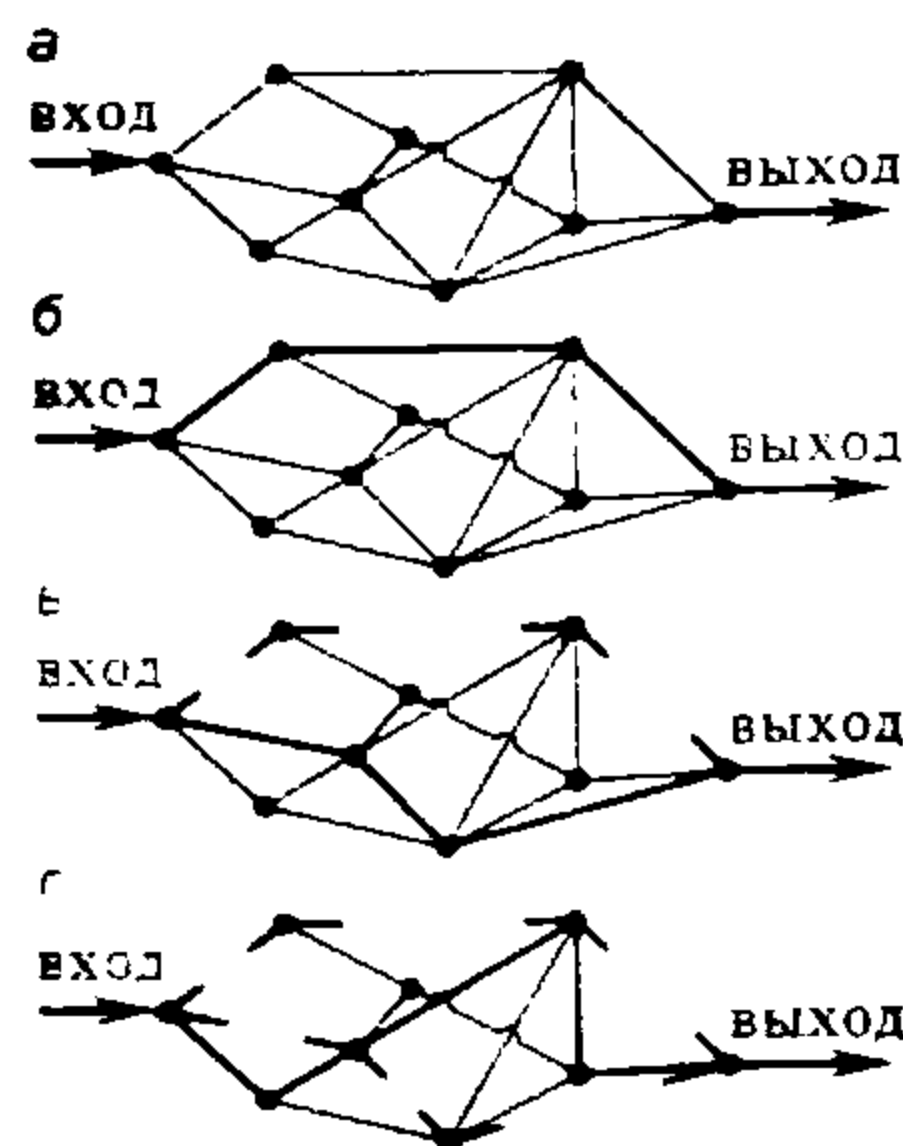


Рис. 3.11

ший условный наикратчайший путь, который оказывается последним из возможных. Заметим, что число независимых путей в каждом подмножестве A_k не может превышать размерности наименьшего минимального разреза графа, т. е. минимального разреза с наименьшим числом входящих в него ребер. В данном случае эта размерность равна трем.

Построенная совокупность независимых путей не является единственной для данного графа. Чтобы построить новую совокупность, исключим из наикратчайшего минимального пути одно из ребер (оно не должно входить в состав наименьшего минимального разреза), после чего начнем опять построение кратчайшего минимального пути в оставшемся подмножестве ребер. Затем процедура продолжается, как и в первой попытке. Пример такого построения приведен на рис. 3.12.

Подобная процедура может быть продолжена: могут быть исключены другие ребра из первого наикратчайшего пути в первом наборе, затем то же самое может быть сделано со вторым набором и т. д. Нужно заметить, что организация построения наборов число их практически может быть небольшим.

Фактически нет необходимости формировать все возможные совокупности независимых минимальных путей. Если нижняя граница Литвака — Ушакова будет построена не на полном множестве всех возможных наборов минимальных путей, то она будет все равно оставаться нижней границей, хотя может оказаться, что и не достигнет ее.

$$P_{\text{Л.У}} = \max_{1 \leq k \leq N} \left\{ P_{\text{Л.У}}^{(k)} \right\} \geq \max_{1 \leq k \leq N} \left\{ P_{\text{Л.У}}^{(k)} \right\}, \quad (3.58)$$

где N — любое натуральное число, меньшее чем N .

Это соображение является очень важным, поскольку позволяет достаточно быстро построить неточную нижнюю границу Литвака — Ушакова.

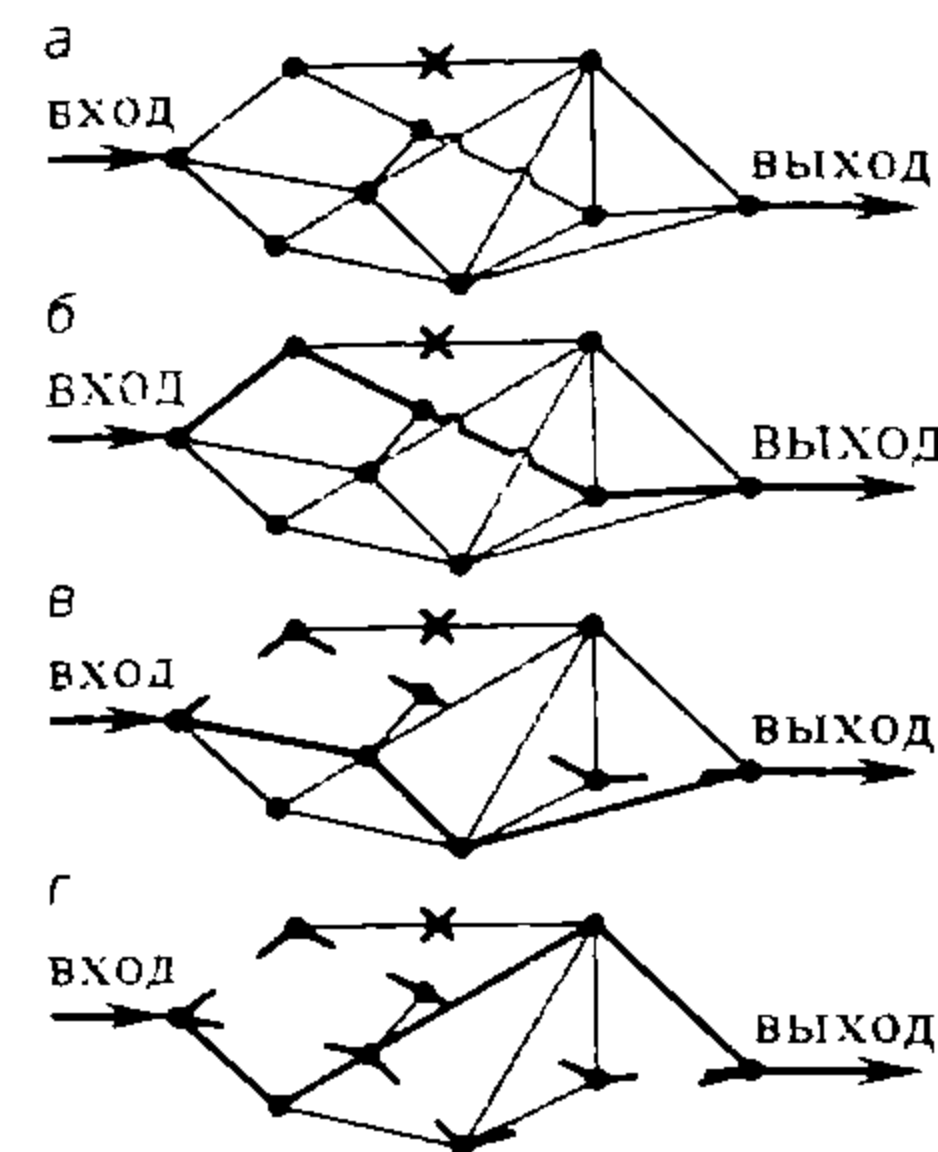


Рис. 3.12

ВЕРХНЯЯ ГРАНИЦА. Перейдем теперь к построению верхней границы Литвака — Ушакова. Как и ранее, структурную функцию j -го минимального разреза обозначим через $\beta_j(X)$. Всего можно выбрать M различных наборов B_1, B_2, \dots, B_M . Для каждого k -го набора независимых разрезов можно записать структурную функцию

$$\varphi_k(X) = \bigcap_{j \in \beta_k} \beta_j(X), \quad (3.59)$$

в которой $\beta_j(X) = \bigcup_{i \in \beta_j} x_i$.

По построению для любого из наборов k выполняется условие $\varphi(X) \leq \varphi_k(X)$, т. е.

$$\varphi(X) \leq \max \{ \varphi_1(X), \varphi_2(X), \dots, \varphi_M(X) \}. \quad (3.60)$$

Применив оператор математического ожидания к обеим частям неравенства (3.60), получаем

$$\bar{P}_{\text{Л.У}} = \max \left\{ E\{\varphi_1(X)\}, E\{\varphi_2(X)\}, \dots, E\{\varphi_M(X)\} \right\}. \quad (3.61)$$

В результате имеем верхнюю границу Литвака — Ушакова:

$$\bar{P}_{\text{Л.У}} = \min \left\{ \prod_{j \in B_1} \left(1 - \prod_{i \in \beta_j(X)} q_i \right), \prod_{j \in B_2} \left(1 - \prod_{i \in \beta_j(X)} q_i \right), \dots, \prod_{j \in B_M} \left(1 - \prod_{i \in \beta_j(X)} q_i \right) \right\}. \quad (3.62)$$

Как и в случае с набором независимых минимальных путей, здесь также можно строить негарантированную нижнюю оценку, используя не все M возможных наборов B_k , а лишь M' — любое натуральное число, меньшее чем M :

$$\bar{P}_{\text{Л.У}} = \min \{ \varphi_1(X), \varphi_2(X), \dots, \varphi_{M'}(X) \} \leq \min \{ \varphi_1(X), \varphi_2(X), \dots, \varphi_M(X) \}. \quad (3.63)$$

В этом случае верхняя граница может ухудшиться.

Таким образом, построены нижняя и верхняя границы Литвака — Ушакова для вероятности связности двухполюсной сети:

$$\begin{aligned} & \max \left\{ 1 - \prod_{\alpha_i \in A_1} \left(1 - \prod_{j \in \alpha_i} p_{ij} \right), 1 - \prod_{\alpha_i \in A_2} \left(1 - \prod_{j \in \alpha_i} p_{ij} \right), \dots, \right. \\ & \left. 1 - \prod_{\alpha_i \in A_N} \left(1 - \prod_{j \in \alpha_i} p_{ij} \right) \right\} \leq E\{\varphi(X)\} \leq \min \left\{ \prod_{j \in B_1} \left(1 - \prod_{i \in \beta_j(X)} q_i \right), \right. \\ & \left. \prod_{j \in B_2} \left(1 - \prod_{i \in \beta_j(X)} q_i \right), \dots, \prod_{j \in B_M} \left(1 - \prod_{i \in \beta_j(X)} q_i \right) \right\} \end{aligned} \quad (3.64)$$

3.4.4. Сравнение граничных оценок Эзари — Прошана и Литвака — Ушакова

Как видно из анализа мостиковой системы, проведенного в предыдущем разделе, верхняя граница Эзари — Прошана дает лучший результат при высоких значениях P , а нижняя граница Литвака — Ушакова — для малых значений p . Заметим, что на практике важнее случай высоких p , т. е. граница Эзари — Прошана в этом смысле оказывается лучшей для высоконадежных систем.

Однако с вычислительной точки зрения границы Эзари — Прошана обладают огромным недостатком: они требуют перечисления всех минимальных путей для вычисления верхней границы и всех минимальных разрезов для вычисления нижней границы. Эта задача сама по себе нетривиальна с вычислительной точки зрения.

Попытка найти приближенную верхнюю или нижнюю границы Эзари — Прошана путем сокращения числа рассматриваемых минимальных путей в первом случае и числа минимальных разрезов во втором может привести к серьезным ошибкам. Представим себе, что, рассматривая верхнюю границу, мы не будем перечислять все минимальные пути, а ограничимся лишь некоторыми из них. Верхняя граница Эзари — Прошана понизится — формально *улучшится*. Продолжим эту процедуру *улучшения* верхней границы Эзари — Прошана. В какой-то момент останутся лишь независимые пути, т. е. верхняя граница Эзари — Прошана совпадет с одной из нижних границ Литвака — Ушакова. Таким образом, полученная «приближенная» оценка Эзари — Прошана потеряет смысл верхней оценки. Это показывает незаконность подобных упрощений: никто не может гарантировать, что на каком-то шаге «улучшения» границы Эзари — Прошана указанным методом она потеряет свой смысл.

То же касается и нижней границы Эзари — Прошана: исключая разрезы из полного набора, можно прийти к тому, что оставшиеся разрезы будут независимыми, и нижняя оценка Эзари — Прошана совпадет с верхней оценкой Литвака — Ушакова. Такая «приближенная» нижняя граница Эзари — Прошана потеряет смысл нижней границы.

Границы Литвака — Ушакова лишены указанных недостатков. Вычисляя их, мы можем остановиться на любом шаге перебора вариантов наборов независимых минимальных путей (или независимых минимальных разрезов). Мы не будем иметь гарантии, что получена наилучшая граница Литвака — Ушакова, однако всегда любая получаемая оценка будет оставаться нижней (или, соответственно, верхней) границей.

§ 3.5. ДЕКОМПОЗИЦИЯ СЕТЕВЫХ СТРУКТУР

3.5.1. Метод рекуррентных вложений

В одной из первых работ, посвященных анализу сложных сетевых систем, Шенноном* были рассмотрены рекуррентные сетевые системы, которые строились по следующему принципу: каждая подсистема верхнего уровня была структурно идентична подсистеме нижнего уровня. Кроме того, полагалось, что все *атомарные* элементы идентичны. Поясним это с помощью достаточно хорошо разобранный ранее мостиковой системы. Пусть имеется исходная мостиковая система. На первом этапе формируем новую вложенную сетевую двухполюсную систему, каждый из элементов которой представляет собой также мостиковую систему (см. рис. 3.13).

Вновь полученную структуру будем называть мостиковой структурой второго порядка. На следующем шаге итеративного построения получается мостиковая структура третьего порядка (см. рис. 3.14) и т. д.

Для систем, построенных таким образом, был предложен простой и элегантный метод вычисления вероятности связности двухполюсных вложенных сетей. Предположим, что *атомарные* элементы системы идентичны и независимы, а вероятность безотказной работы каждого из них равна p . Обозначим функцию связности для мостиковой системы через $p_1 = h(p)$. Тогда для системы на рис. 3.13 можно записать функцию связности в виде

$$p_2 = h(p_1) = h(h(p)), \quad (3.65)$$

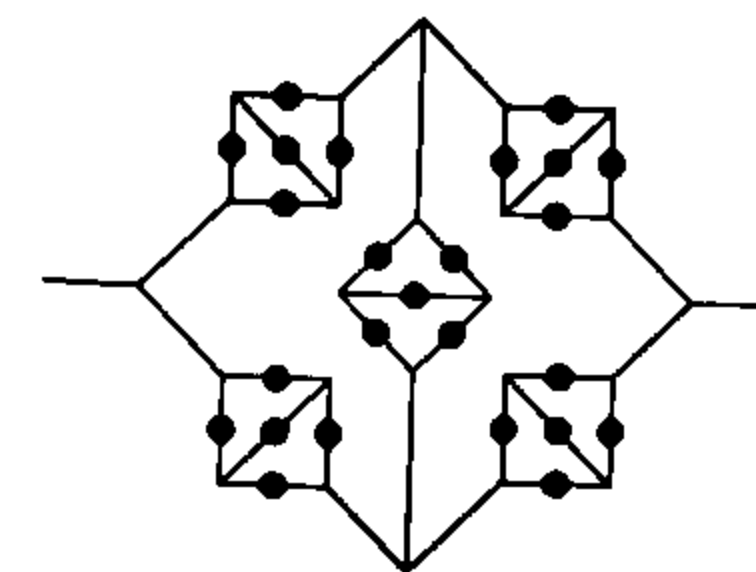


Рис. 3.13

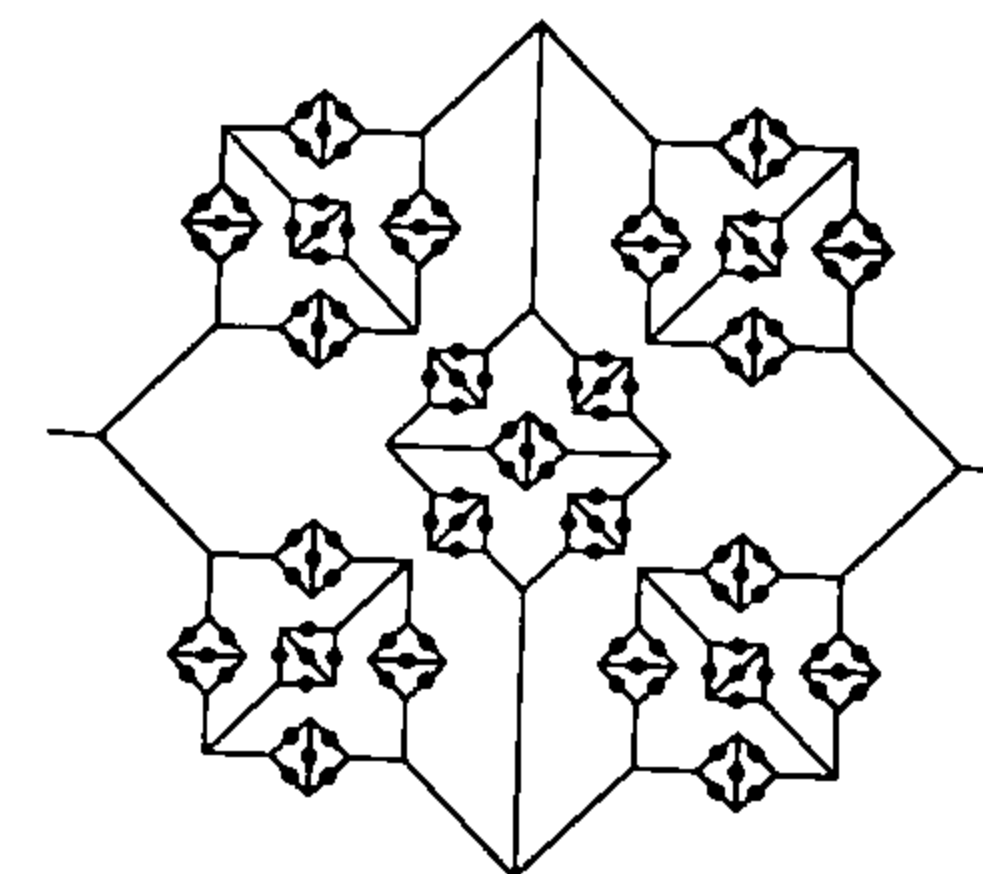


Рис. 3.14

* К. Э. Шеннон (Shannon Cl.E., 1916—2001) — американский математик и инженер, один из создателей математической теории информации.

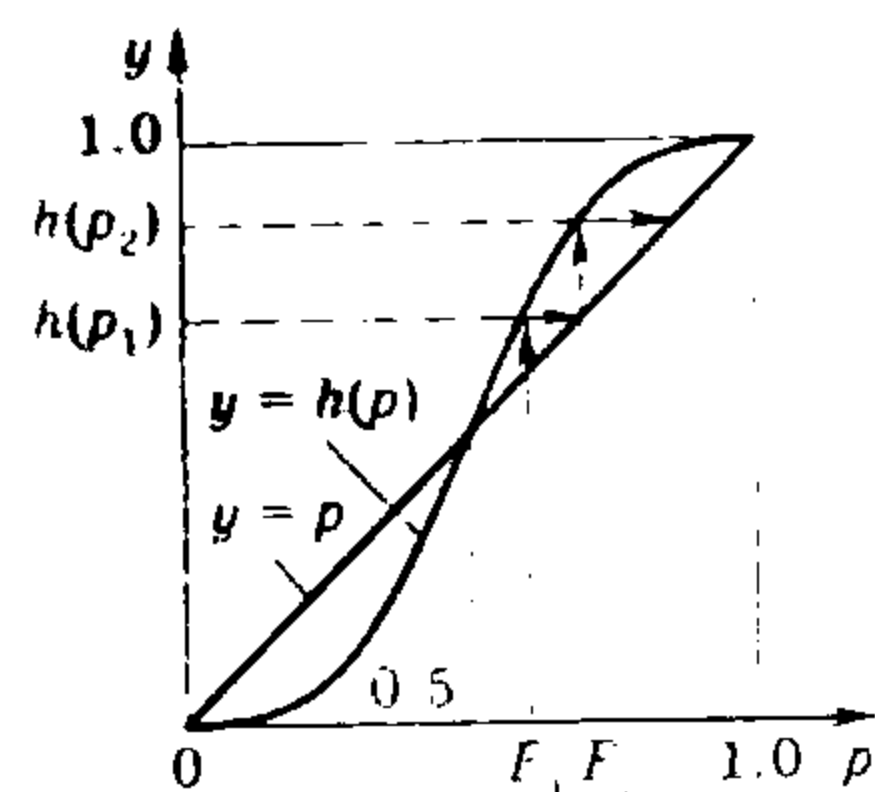


Рис. 3.15

а для системы на рис. 3.14

$$p_3 = h(p_2) = h(h(h(p))) \quad (3.66)$$

и т. д.

Структура вычислительного процесса проиллюстрирована рис. 3.15, где изображена зависимость $h(p)$. Действительно, надежность элемента p_1 соответствует значению $h(p_1) = p_2$, элемента p_2 — значению $h(p_2) = p_3$ и т. д.

Вложенная сетевая система может быть построена и с использованием иных исходных структур.

Следует отметить, что подобного рода красивые примеры фактически не встречаются в инженерной практике.

3.5.2. Метод разбиения на модули

Метод заключается в выделении из сети модулей, которые по своей структуре являются достаточно простыми. Пояснить этот метод лучше всего с помощью примера.

Рассмотрим сеть, представленную на рис. 3.16. Здесь через S обозначена система в целом, а через S_i и S_{ij} — подсистемы нижних уровней.

Обозначим через $P(S)$ вероятность связности графа, а через $P(S_i)$ и $P(S_{ij})$ — соответствующие вероятности для подсистем различного уровня. Используем в дальнейшем обозначение $Q = 1 - P$. Для рассматриваемой сетевой системы после разложения на модули первого уровня можно записать:

$$P(S) = P(S_1)[1 - Q(S_2)][1 - Q(S_3)], \quad (3.67)$$

а в свою очередь, $P(S_1) = P(S_{11})P(S_{12})$, $P(S_2) = 1 - Q(S_{21})Q(S_{22})$ и $P(S_3) = P(S_{31})P(S_{32})$.

Видно, что, кроме S_{32} , являются приводимыми, а S_{32} представляет собой уже известную мостиковую схему.

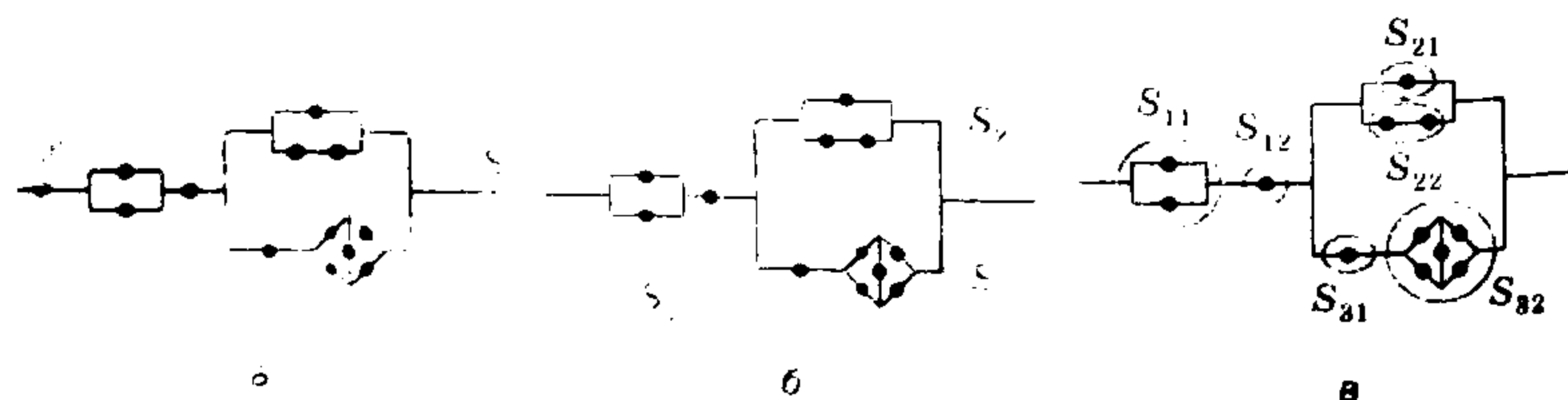


Рис. 3.16

Возможно редуцировать модули не до уровня приводимых систем, а выделив некий модуль, затем применять граничные оценки Эзари — Прошана или Литвака — Ушакова. Например, для предыдущей сетевой системы оставившись на уровне модулей S_1 , S_2 и S_3 , вести дальнейший расчет верхней граничной и верхней оценки соответственно по формулам

$$\bar{P}(S) = \bar{P}(S_1)[1 - \bar{Q}(S_2)\bar{Q}(S_3)], \quad (3.68)$$

$$\bar{P}(S) = \bar{P}(S_1)[1 - \underline{Q}(S_2)\underline{Q}(S_3)]. \quad (3.69)$$

Здесь в нижнюю границу \underline{P} входят верхние оценки $\bar{Q}(S_2)$ и $\bar{Q}(S_3)$, а в верхнюю границу \bar{P} — нижние границы $\underline{Q}(S_2)$ и $\underline{Q}(S_3)$.

При всей заманчивости такого подхода не всегда удастся так просто найти подходящие для последующего анализа модули. Еще большие сложности возникают при анализе сетей, описываемых непланарными графами.

3.5.3. Модифицированный метод Литвака — Ушакова

В основе метода лежит идея построения граничных оценок Литвака — Ушакова для вероятности связности сетевой системы. Ради удобства представления проведем изложение с помощью примера.

Рассмотрим сетевую структуру, представленную на рис. 3.17, а. Нижняя граница вероятности связности находится путем выделения некоторого минимального пути и включения его параллельно оставшейся сети. Выделим путь, проходящий через верхние вершины, и исключим его из графа. Оставшийся подграф и параллельный ему выделенный путь представлены на рис. 3.17, б, а на рис. 3.17, в — вновь образованный граф в более наглядной форме. Заметим, что в графе появляются *расщепленные* вершины, поскольку исключение ребер какого-либо пути не мешает использовать принадлежащие ему вершины для других путей. Сам по себе подграф, параллельный выделенному пути, не столь уж прост для численного анализа, поэтому при необходимости он может быть еще более упрощен. Когда будут выделены все минимальные пути, такая оценка совпадет с одной из нижних границ Литвака — Ушакова.

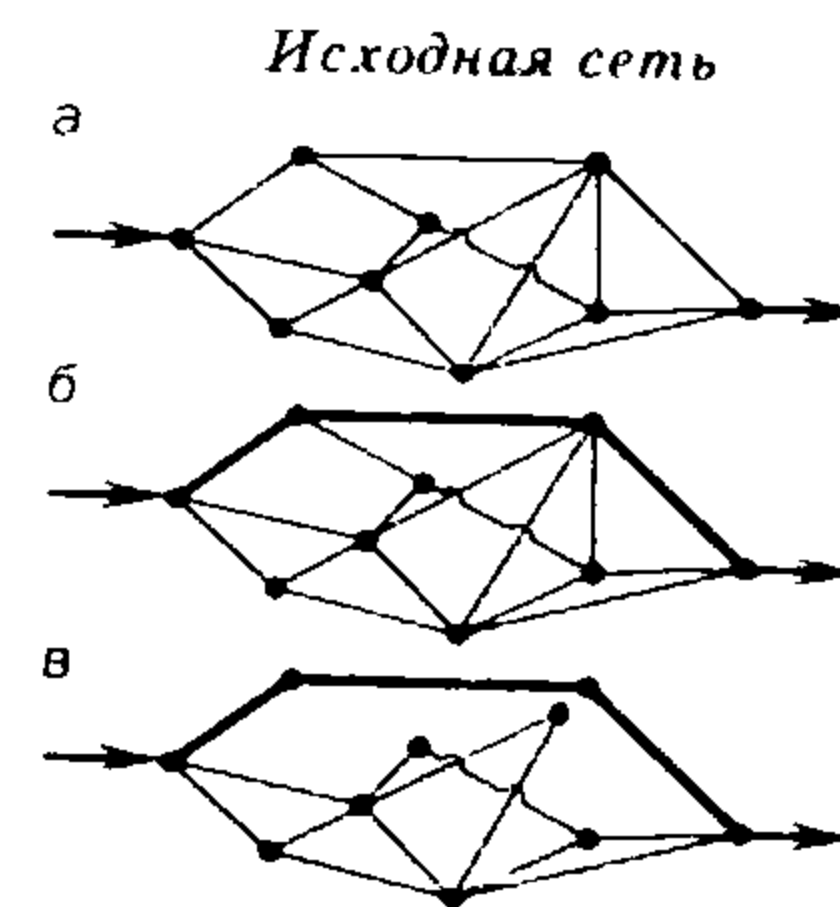


Рис. 3.17

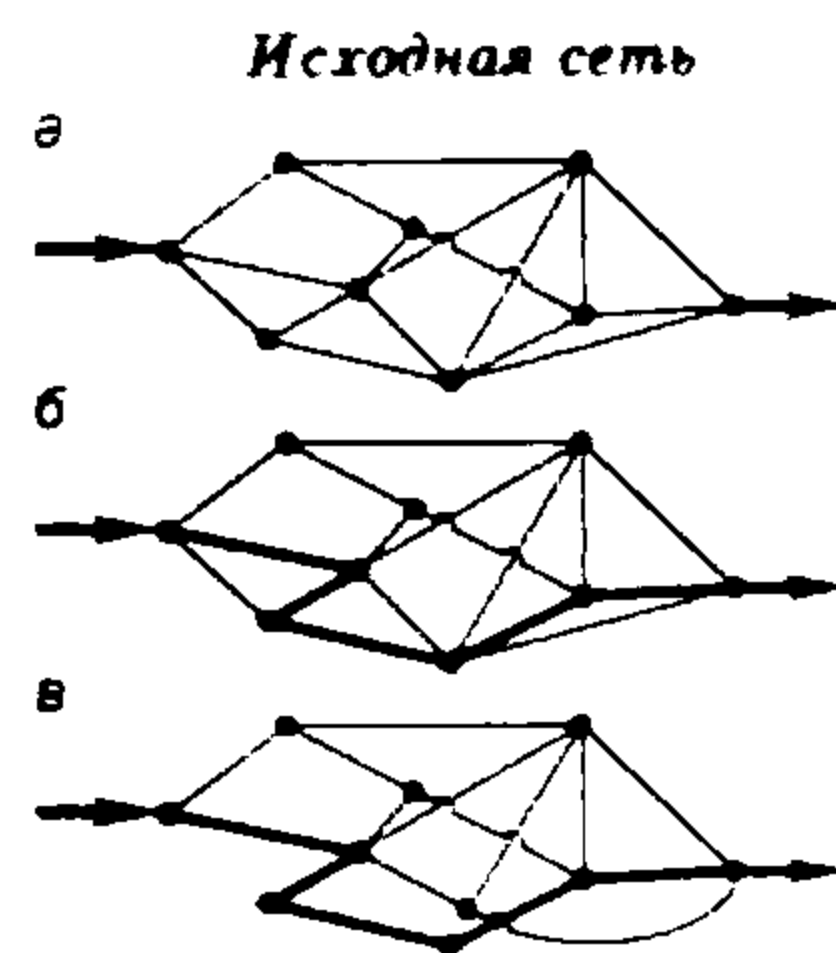


Рис 3.18

Подобное разложение не является единственным. Пример другого разложения того же графа относительно другого выбранного минимального пути приведен на рис. 3.18.

Число возможных итераций в каждой реализации процесса декомпозиции не превышает числа ребер в самом *тонком* минимальном разрезе (условие независимости путей).

При предложенной декомпозиции сетевой структуры могут образовываться или *висячие* ребра (т. е. ребра, подсоединенные лишь к одной вершине), или даже изолированные ребра. Это лишь раз подчеркивает то утверждение, что данная процедура приводит к нижней оценке (некоторые из элементов системы фактически заменяются на абсолютно ненадежные).

ВЕРХНЯЯ ГРАНИЦА. Верхняя граница вероятности связности находится путем выделения некоторого минимального разреза и включения его последовательно к оставшейся подсети. Рассмотрим прежнюю сетевую структуру (рис. 3.19, а).

Выделим в рассматриваемом графе минимальный разрез, замыкающий к левому полюсу сети (рис. 3.19, б). Результирующий граф приведен на рис. 3.19, в.

Подобное разложение не является единственным. На рис. 3.20 выделен разрез в центре графа, разбивающий его на два подграфа — слева и справа от выбранного разреза.

Левые вершины ребер выбранного разреза стягиваются в одну левую вершину, а все правые вершины стягиваются в одну пра-

вую вершину. Процесс стягивания левых вершин разреза в одну точку эквивалентен тому, что соответствующие ребра между исходными левыми вершинами полагаются абсолютно надежными. То же относится и к процессу стягивания в одну точку правых вершин того же разреза. Поэтому полученная оценка вероятности связности будет верхней.

Естественно, что качество оценки определяется успешностью выбора исходного разреза. При необходимости процедуру разложения оставшегося графа можно продолжить. Число таких итераций не превышает числа ребер в кратчайшем минимальном пути. В конечном счете, после того как будут выделены все минимальные разрезy, эта оценка совпадет с одной из верхних границ Литвака — Ушакова.

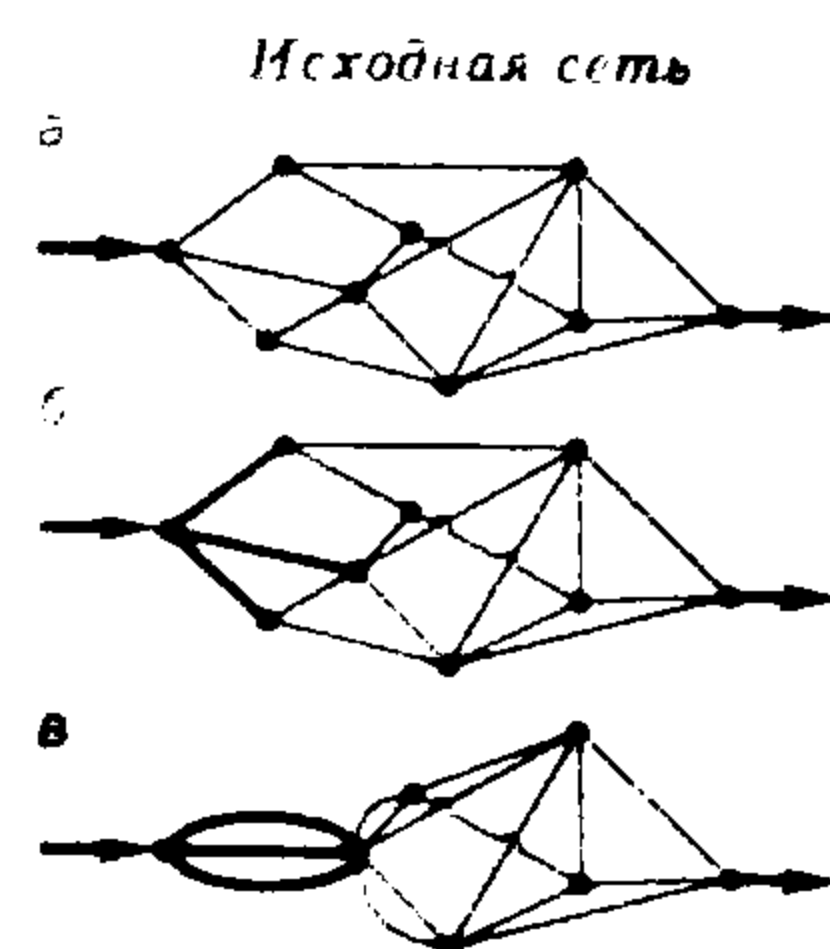


Рис. 3.19

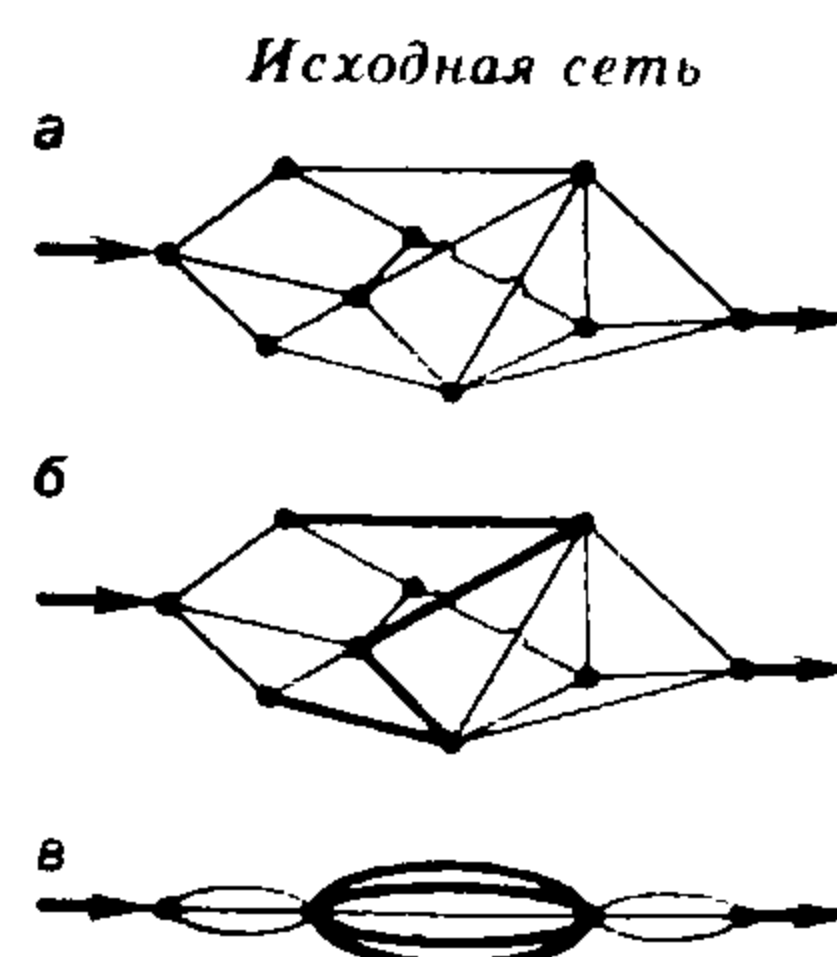


Рис 3.20

Глава 4

Марковские модели восстанавливаемых систем

Во многих задачах надежности предполагается, что распределения времени до отказа и времени восстановления являются экспоненциальными. Такой подход показал свою применимость на практике, поскольку обычно бывают известны только средние значения времени безотказной работы и времени восстановления, что предопределяет использование однопараметрического распределения. Кроме того, в предположении экспоненциальности распределений удается многие задачи надежности решить аналитически.

§ 4.1. ПУАССОНОВСКИЙ ПОТОК

4.1.1. Характеристические свойства

Пуассоновским потоком называется случайный точечный поток, расстояния между событиями в котором имеют экспоненциальное распределение.

В качестве событий в дальнейшем будем рассматривать отказы, а сам поток будем называть пуассоновским потоком отказов. У такого потока число отказов k , произошедших в некотором фиксированном интервале $[0, t]$ подчиняется распределению Пуассона:

$$P_k(t) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t). \quad (4.1)$$

Величина λ называется интенсивностью пуассоновского потока — она характеризует среднее число отказов, происходящих в единицу времени.

Отсюда следует, что $\lambda = 1/T$, где T есть среднее расстояние между событиями данного потока. Поток Пуассона определяется следующими свойствами.

- Стационарность, т. е. инвариантность по отношению к операции сдвига. Иными словами, характеристики потока (например, среднее число отказов) на интервале длины t зависят только от его длины.
- Отсутствие последствия, или марковское свойство, т. е. независимость будущего развития потока от всей его предыстории.
- Ординарность, т. е. отсутствие точек концентрации или же одновременного возникновения нескольких отказов. Иначе говоря, при стремлении длины интервала к нулю вероятность появления более одного отказа имеет более высокий порядок малости по сравнению с вероятностью появления ровно одного отказа.

4.1.2. Метод предельного перехода

Из трех сформулированных почти что качественных свойств потока Пуассона следует строгое формальное его определение, выражаемое формулой (4.1). Действительно, рассмотрим некоторый произвольный интервал времени, скажем, t . Разобьем этот интервал на малые интервалы Δ_i так, чтобы

$$\sum_{i=1}^n \Delta_i = t.$$

Среднее число отказов (математическое ожидание) за время Δ_i по определению равно

$$E(\Delta_i) = \lambda \Delta_i. \quad (4.2)$$

С другой стороны, для того же интервала Δ_i можно записать и другое определение этого среднего:

$$E(\Delta_i) = \sum_{j=0}^{\infty} j p_j(\Delta_i), \quad (4.3)$$

здесь $p_j(\Delta_i)$ есть вероятность того, что на интервале Δ_i произойдет ровно j отказов. С учетом свойства ординарности можно записать условие полной группы событий:

$$p_0(\Delta_k) + p_1(\Delta_k) + o(\Delta_k) = 1, \quad (4.4)$$

где $o(\Delta_k)$ означает величину более высокого порядка малости, чем $p_1(\Delta_k)$. Используя свойство ординарности, выражение (4.3) можно переписать в виде

$$\lim_{\Delta_k \rightarrow 0} E(\Delta_k) = p_1(\Delta_k) \quad (4.5)$$

или с учетом (4.2) и (4.4):

$$p_1(\Delta_i) = \lambda \Delta_i \quad (4.6)$$

и

$$p_0(\Delta_i) = 1 - \lambda \Delta_i. \quad (4.7)$$

Проанализируем, чему равна вероятность того, что на интервале t не произойдет ни одного отказа. В силу отсутствия последующих действий можно записать:

$$P(t) = \prod_{i=1}^n p_0(\Delta_i). \quad (4.8)$$

В пределе при равномерном стремлении Δ_i к нулю и условии, что $\sum \Delta_i = t$, получаем

$$P(t) = \lim_{\Delta \rightarrow 0} (1 - \lambda \Delta)^{t/\Delta} = \exp(-\lambda t). \quad (4.9)$$

Это и завершает доказательство того, что рассматриваемый поток отказов является пуассоновским.

4.1.3. Метод преобразований Лапласа

Применим теперь к решению той же задачи метод преобразований Лапласа, более формальный, но и более компактный; далее он будет использоваться при решении других задач теории надежности.

Преобразование Лапласа функции $y(t)$ определяется как

$$\varphi(s) = \int_0^{\infty} y(t) e^{-st} dt. \quad (4.10)$$

Взаимно-однозначное соответствие функции-оригинала и ее преобразования будем обозначать в виде $y(t) \rightleftharpoons \varphi(s)$.

Преобразование Лапласа оказывается крайне полезным при решении дифференциальных уравнений, поскольку вместо исходной системы мы решаем систему алгебраических уравнений в преобразованиях Лапласа, и затем через обратное преобразование Лапласа приходим к искомому решению. Такой путь решения требует трех простых шагов вместо одного, но сложного; процедура поясняется рис. 4.1. В табл. 4.1 приводится краткая сводка основных преобразований Лапласа, используемых в книге.

Заметим, что если взять преобразование Лапласа $\varphi_F(s)$ от функции $P(t) = 1 - F(t)$, то можно найти значение среднего для распределения $F(t)$ как

$$\varphi(s)|_{s=0} = \int_0^{\infty} P(t) e^{-st} dt|_{s=0} = \int_0^{\infty} P(t) dt = M_1. \quad (4.11)$$

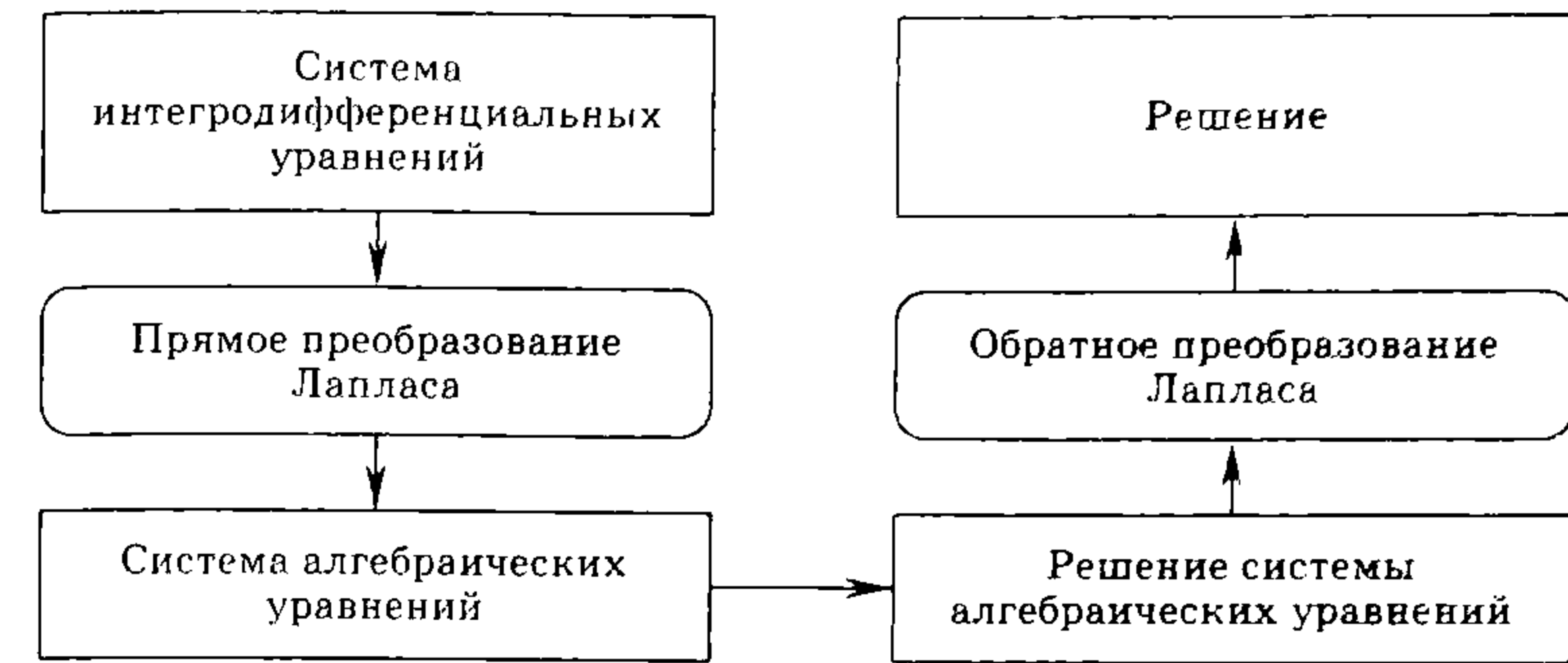


Рис. 4.1

Для вероятности появления k отказов на интервале t можно записать формулу полной вероятности с учетом характеристических свойств пуассоновского потока в виде

$$P_k(t + \Delta) = P_k(t)P_0(\Delta) + P_{k-1}(t)P_1(\Delta) + o(\Delta), \quad (4.12)$$

где $P_k(t)$ есть вероятность того, что система в момент времени t находится в состоянии k . Из (4.12) получаем

$$\frac{P_k(t + \Delta) - P_k(t)}{\Delta} = -\lambda P_k(t) + \lambda P_{k-1}(t) + o(\Delta). \quad (4.13)$$

Таблица 4.1

Оригинал	Изображение
$F(t)$	$\varphi(s)$
$\frac{d}{dt} F(t)$	$s\varphi(s)$
$\int_0^t F(x) dx$	$\frac{\varphi(s)}{s}$
$aF_1(t) + bF_2(t)$	$a\varphi_1(s) + b\varphi_2(s)$
$F \circ G(t) = \int_0^t F(x)G(t-x)dx$	$\varphi_F(s)\varphi_G(s)$
$\exp(-\lambda t)$	$\frac{1}{\lambda + s}$
$\frac{(\lambda t)^k}{k!} \exp(-\lambda t)$	$\frac{\lambda^k}{(\lambda + s)^{k+1}}$
C	C/s

После предельного перехода при $\Delta \rightarrow 0$ равенство (4.13) дает дифференциальное уравнение вида

$$\frac{d}{dt} P_k(t) = -\lambda P_k(t) + \lambda P_{k-1}(t). \quad (4.14)$$

Обозначим преобразование Лапласа от функции $P_k(t)$ через $\varphi_k(s)$.

Тогда дифференциальное уравнение (4.14) можно записать в преобразованиях Лапласа как

$$s\varphi_k(s) = -\lambda\varphi_k(s) + \lambda\varphi_{k-1}(s), \quad (4.15)$$

или

$$\varphi_k(s) = \frac{\lambda\varphi_{k-1}(s)}{\lambda + s}. \quad (4.16)$$

Будем решать это уравнение рекуррентным методом, начав с вероятности $P_0(t)$. На основании (4.14) для $k = 0$ имеем

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t). \quad (4.17)$$

Решение этого дифференциального уравнения требует еще выбора начального условия, которое в данном случае имеет естественный вид $P_0(0) = 1$, т. е. система в начальный момент функционирования находится в состоянии « \circ » без отказавших элементов.

Преобразование Лапласа для дифференциального уравнения (4.17) можно записать в виде

$$s\varphi_0(s) - 1 = -\lambda\varphi_0(s), \quad (4.18)$$

откуда следует

$$\varphi_0(s) = \frac{1}{\lambda + s}. \quad (4.19)$$

Функцией-оригиналом преобразования (4.19) является $\exp(-\lambda t)$. Из (4.19), систематически используя (4.16), находим

$$\varphi_k(s) = \frac{\lambda^k}{(\lambda + s)^{k+1}}. \quad (4.20)$$

Оригиналом этого изображения является

$$\frac{(\lambda t)^k}{k!} \exp(-\lambda t). \quad (4.21)$$

Таким образом, поток, определяемый указанными характеристиками, имеет пуассоновское распределение числа отказов в течение времени t с интенсивностью λ (или с длительностью $1/\lambda$).

§ 4.2. ВОССТАНАВЛИВАЕМЫЙ ЭЛЕМЕНТ

Нахождение различных показателей надежности восстанавливаемого элемента будет изложено весьма детально с тем, чтобы избежать аналогичных, но гораздо более громоздких выкладок в более сложных случаях.

4.2.1. Нестационарный коэффициент готовности

Нестационарный КГ определяется как вероятность того, что в заданный момент времени t элемент будет в состоянии работоспособности при известном состоянии в начальный момент времени.

Предполагается, что элемент может быть в двух состояниях: работоспособности и отказа. Граф переходов для элемента представлен на рис. 4.2.

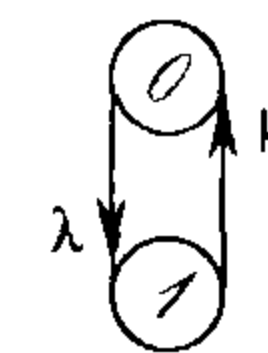


Рис. 4.2

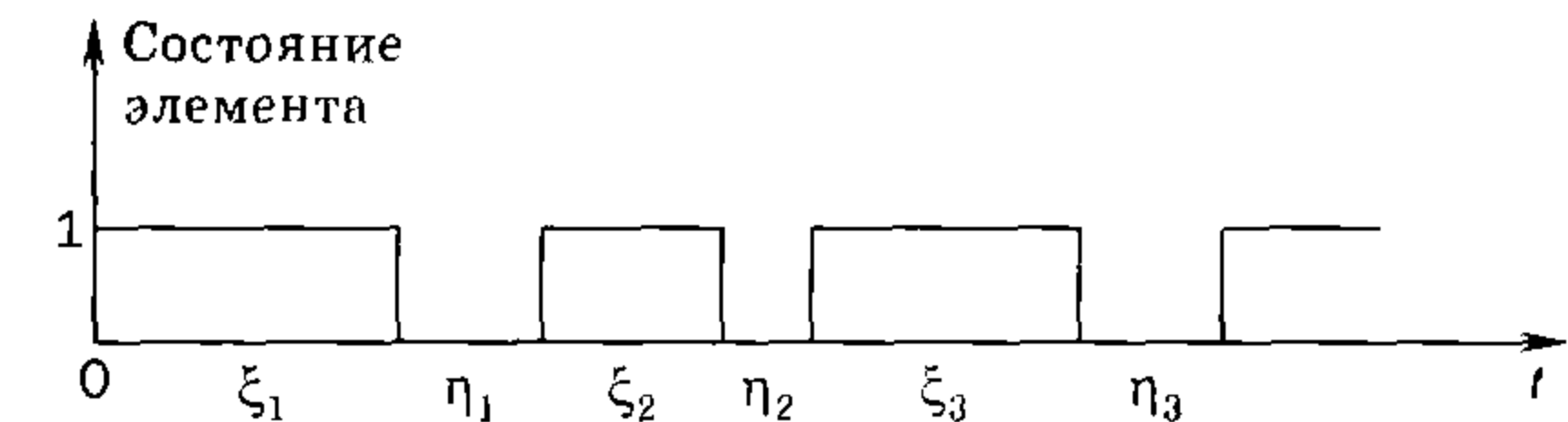


Рис. 4.3

Переход из состояния работоспособности в состояние отказа происходит с интенсивностью $\lambda = 1/T$, а переход из состояния отказа в состояние работоспособности — с интенсивностью $\mu = 1/\tau$, где T есть среднее время работы до отказа элемента, а τ — среднее время его восстановления. После каждого восстановления элемент предполагается полностью обновленным, т. е. его вероятностные характеристики становятся такими же, как и у совершенно нового элемента.

Процесс функционирования восстанавливаемого элемента описывается альтернирующим процессом восстановления, представляющим собой чередующиеся взаимно независимые случайные интервалы работоспособности ξ и интервалы восстановления η (рис. 4.3).

На основании графа переходов (рис. 4.2) нетрудно записать дифференциальное уравнение динамического баланса для вероятностей (см. Приложение 3):

$$\begin{cases} \frac{d}{dt} P_0(t) = -\lambda P_0(t) + \mu P_1(t), \\ P_0(t) + P_1(t) = 1. \end{cases} \quad (4.22)$$

В качестве начальных условий возьмем для общего случая $P_0(0) = p$ и $P_1(0) = q$, т. е. предположим, что элемент может с вероятностью p быть в работоспособном состоянии, а с вероятностью $q = (1 - p)$ — в состоянии отказа. Для системы уравнений (4.22) можно записать систему алгебраических уравнений в преобразованиях Лапласа:

$$\begin{cases} -p + s\varphi_0(s) = -\lambda\varphi_0(s) + \mu\varphi_1(s), \\ \varphi_0(s) + \varphi_1(s) = \frac{1}{s}, \end{cases} \quad (4.23)$$

или в канонической форме

$$\begin{bmatrix} \lambda + s & -\mu \\ s & s \end{bmatrix} \begin{bmatrix} \varphi_0(s) \\ \varphi_1(s) \end{bmatrix} = \begin{bmatrix} p \\ 1 \end{bmatrix}. \quad (4.24)$$

Применим правило Крамера к решению алгебраической системы уравнений (4.24):

$$\varphi_0(s) = \frac{\begin{vmatrix} p & -\mu \\ 1 & s \end{vmatrix}}{\begin{vmatrix} \lambda + s & -\mu \\ s & s \end{vmatrix}} = \frac{ps + \mu}{s^2 + (\lambda + \mu)s}. \quad (4.25)$$

Для обращения полученного преобразования Лапласа представим правую часть (4.25) в виде суммы слагаемых типа $a/(s + b)$, где, вообще говоря, b может быть равным нулю. Знаменатель в соотношении (4.25) может быть представлен в виде

$$s^2 + (\lambda + \mu)s = (s - s_1)(s - s_2),$$

где s_1 и s_2 — корни полинома: $s_1 = 0$, $s_2 = -(\lambda + \mu)$. Для преобразования Лапласа (4.25) можно записать эквивалентную формулу

$$\varphi_0(s) = \frac{A}{s - s_1} + \frac{B}{s - s_2} = \frac{A}{s} + \frac{B}{s + \lambda + \mu}. \quad (4.26)$$

В силу эквивалентности (4.25) и (4.26) можно записать:

$$\frac{ps + \mu}{s(s + \lambda + \mu)} = \frac{A}{s} + \frac{B}{s + \lambda + \mu} = \frac{(A + B)s + A(\lambda + \mu)}{s(s + \lambda + \mu)}. \quad (4.27)$$

Итак, A и B — неизвестные константы, которые находятся из условия эквивалентности полиномов в числителях (4.25) и (4.27): полиномы равны, если равны коэффициенты при аргументах с соответствующими степенями. Это позволяет написать новую систему алгебраических уравнений для неизвестных A и B :

$$\begin{cases} A + B = p, \\ A(\lambda + \mu) = \mu. \end{cases} \quad (4.28)$$

Отсюда легко находятся

$$A = \frac{\mu}{\lambda + \mu}, \quad B = p - \frac{\mu}{\lambda + \mu}. \quad (4.29)$$

Таким образом, интересующее нас преобразование Лапласа запишется в виде

$$\varphi_0(s) = \frac{\mu}{\lambda + \mu} \frac{1}{s} + \frac{\lambda p - \mu q}{\lambda + \mu} \frac{1}{\lambda + \mu + s}. \quad (4.30)$$

В результате обращения этого преобразования Лапласа получаем выражение для нестационарного КГ:

$$K(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda p - \mu q}{\lambda + \mu} \exp[-(\lambda + \mu)t]. \quad (4.31)$$

Во многих практических случаях рассматривается начальный период работы объекта, причем предполагается, что в момент $t = 0$ он работоспособен, т. е. начальное условие имеет вид $P_0(0) = 1$. В этом случае (4.31) преобразуется в более простое выражение

$$K(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t]. \quad (4.32)$$

Эта функция представлена на рис. 4.4.

4.2.2. Стационарный коэффициент готовности

Если $t \rightarrow \infty$, то $K(t)$ сходится к некоторому *стационарному* значению, как это видно и из рис. 4.4. Действительно, при $t \rightarrow \infty$ из (4.32) имеем

$$K = \lim_{t \rightarrow \infty} K(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \lim_{t \rightarrow \infty} \exp[-(\lambda + \mu)t],$$

или, окончательно,

$$K = \frac{\mu}{\lambda + \mu} = \frac{T}{T + \tau}, \quad (4.33)$$

где T есть среднее время работы между отказами, а τ есть среднее время простоя (восстановления работоспособности). Стационарный КГ не зависит от начальных условий.

Вообще говоря, коэффициент готовности может быть определен и иначе, а именно как средняя доля времени, в течение которого объект находится в состоянии работоспособности.

Средняя длина цикла равна $T + \tau$, а среднее время работоспособного состояния на нем есть T , т. е. путем простых рассуждений приходим к той же формуле (4.33).

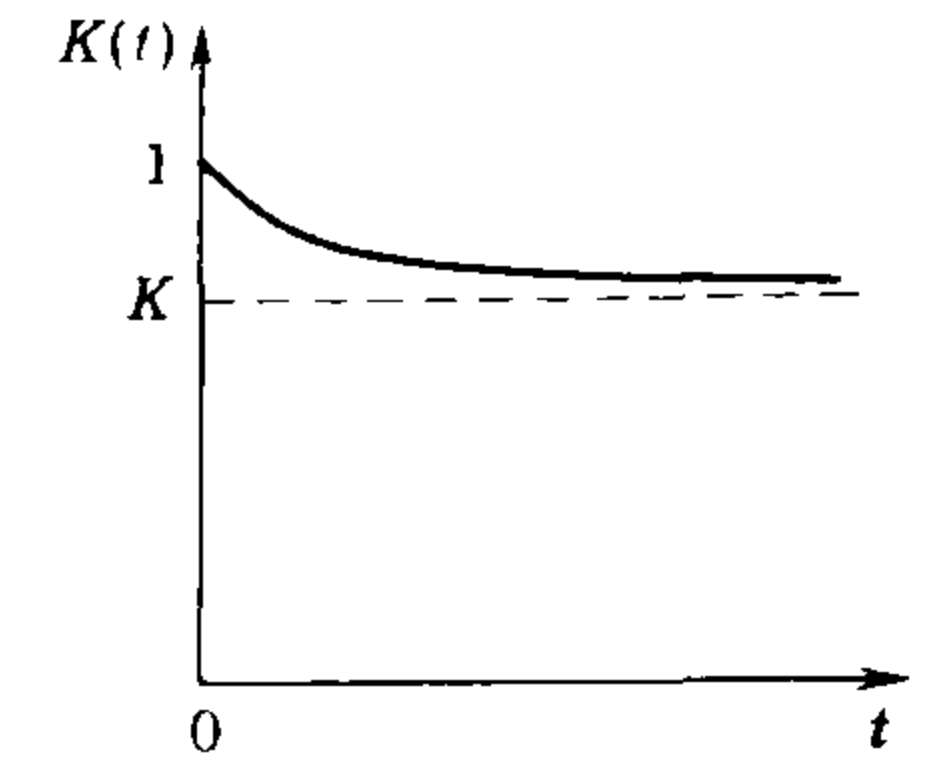


Рис. 4.4

Получать значение стационарного КГ, или просто коэффициента готовности, упрощая выражение для нестационарного КГ, не является лучшим способом решения. В подавляющем большинстве случаев (особенно для сложных систем) удобнее находить коэффициент готовности непосредственно из алгебраической системы уравнений баланса, описывающих стационарное поведение системы. В данном случае на основании рис. 4.2 можно сразу же записать систему уравнений, включающую уравнение баланса для разреза между состояниями 0 и 1: $\lambda P_0 = \mu P_1$. Добавив условие нормировки, получаем следующую систему уравнений:

$$\begin{cases} \lambda P_0 - \mu P_1 = 0, \\ P_0 + P_1 = 1. \end{cases} \quad (4.34)$$

Теперь решение находится сразу же:

$$K = P_0 = \frac{1}{1 + \lambda/\mu} = \frac{\mu}{\lambda + \mu} = \frac{T}{T + \tau}. \quad (4.35)$$

4.2.3. Вероятность безотказной работы

Этот случай соответствует уже рассмотренному ранее случаю невосстанавливаемого элемента с экспоненциальным временем работы до отказа. Ради методологического единообразия изложения получим этот элементарный результат стандартным для данной главы путем. Рассмотрим простейший нетранзитивный* граф переходов (рис. 4.5). Процесс начинается в состоянии 0 и через некоторое время завершается переходом в поглощающее состояние 1, откуда он уже никогда не возвращается.

Избегая тривиальных объяснений, запишем дифференциальное уравнение:

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t). \quad (4.36)$$

Применив преобразование Лапласа и приняв начальное условие $P_0(0) = 1$, получим

$$-1 + s\varphi_0(s) = -\lambda\varphi_0(s), \quad (4.37)$$

или

$$\varphi_0(s) = \frac{1}{\lambda + s}, \quad (4.38)$$



Рис. 4.5

* Напомним, что граф называется транзитивным, если за конечное число шагов можно попасть из произвольной начальной вершины в любую заданную вершину. В противном случае граф называется нетранзитивным. В частности, все графы с поглощающими состояниями являются нетранзитивными.

что соответствует оригиналу

$$P_0(t) = \exp(-\lambda t). \quad (4.39)$$

Вероятность безотказной работы восстанавливаемого элемента совпадает с аналогичным показателем надежности невосстанавливаемого элемента.

4.2.4. Среднее время работы до отказа

Этот показатель может быть в данном случае легко найден путем интегрирования (4.39):

$$T = \int_0^{\infty} \exp(-\lambda t) dt = \frac{1}{\lambda}, \quad (4.40)$$

или же путем подстановки $s = 0$ в (4.38):

$$T = \frac{1}{\lambda + s} \Big|_{s=0} = \frac{1}{\lambda}. \quad (4.41)$$

4.2.5. Среднее время восстановления

Находя этот показатель, нужно обратиться к графу переходов, где поглощающее состояние есть состояние 0, а процесс начинается из состояния 1 (см. рис. 4.6).

Результат можно получить, используя то обстоятельство, что такой случай является «зеркальным» по отношению к предыдущему:

$$\tau = \frac{1}{\mu}. \quad (4.42)$$

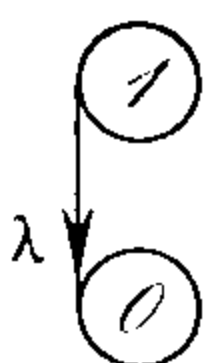


Рис. 4.6

§ 4.3. ПОСЛЕДОВАТЕЛЬНАЯ СИСТЕМА

4.3.1. Описание системы

Рассматривая последовательную восстанавливаемую систему из n элементов, мы сталкиваемся с несколькими ситуациями.

- Во время замены отказавшего элемента остальные элементы либо продолжают находиться в рабочем режиме, либо вся система отключается, т. е. новый отказ за время восстановления отказавшего элемента возникнуть не может.

- Само восстановление может быть неограниченным, когда возможно одновременное восстановление любого числа отказавших элементов, либо ограниченным, когда одновременно может осуществляться восстановление не более чем r , $r < n$, отказавших элементов. В этом случае элементы зависимы по восстановлению.

4.3.2 Система с отключением при отказе

Некоторые системы полностью отключаются на время восстановления отказавшего элемента. Такая ситуация возникает, если процесс ремонта отказавшего элемента может быть связан с определенным риском для обслуживающего персонала (например, восстановление высоковольтных линий). В этом случае все элементы отключаются, и их отказ во время текущего ремонта предполагается невозможным. Элементы системы могут быть различными и по интенсивности отказов λ_i , и по интенсивности восстановления μ_i . Граф переходов, описывающий эту систему, представлен на рис. 4.7.

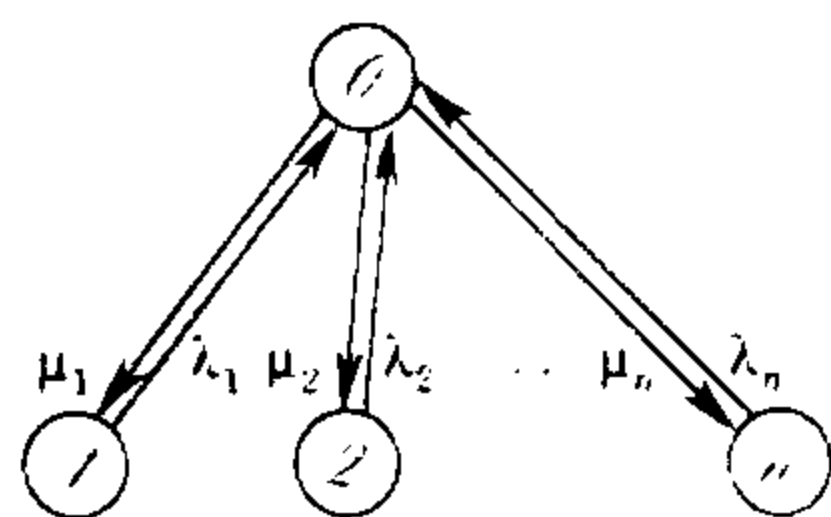


Рис. 4.7

На этом графе состояние 0 означает состояние полной исправности системы, а состояние i , $i = 1, 2, \dots, n$ — отказ i -го элемента системы.

ВЕРОЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ системы определяется как и в случае невозстанавливаемой системы:

$$P_{\text{сист}} = \exp\left(-t_0 \sum_{i=1}^n \lambda_i\right). \quad (4.43)$$

Для высоконадежных систем, у которых $\max \lambda_i t_0 \ll 1/n$, можно записать приближение

$$P_{\text{сист}} \approx 1 - t_0 \sum_{i=1}^n \lambda_i. \quad (4.44)$$

СРЕДНЕЕ ВРЕМЯ БЕЗОТКАЗНОЙ РАБОТЫ. Из (4.43), интегрируя, находим

$$T_{\text{сист}} = \left(\sum_{i=1}^n \lambda_i \right)^{-1}. \quad (4.45)$$

СРЕДНЕЕ ВРЕМЯ ВОССТАНОВЛЕНИЯ находится как средневзвешенное. Действительно, с вероятностью $\lambda_i / \sum_{i=1}^n \lambda_i$ в системе происходит отказ элемента i -го типа, при этом длительность простоя системы равна τ_i . Таким образом, можно записать:

$$\tau_{\text{сист}} = \sum_{i=1}^n \tau_i \frac{\lambda_i}{\Lambda_{\text{сист}}} = \frac{\sum_{i=1}^n \lambda_i \tau_i}{\sum_{i=1}^n \lambda_i}, \quad (4.46)$$

где $\Lambda_{\text{сист}} = \sum_{i=1}^n \lambda_i$.

Заметим, что время восстановления имеет гиперэкспоненциальное распределение (см. Приложение 2, с. 209):

$$Pr\{\eta_{\text{сист}} \geq t\} = \frac{\lambda_i \exp(-\lambda_i t)}{\sum_{i=1}^n \lambda_i}. \quad (4.47)$$

Это означает, что оно имеет убывающую функцию интенсивности, т. е. является *молодеющим* распределением.

КОЭФФИЦИЕНТ ГОТОВНОСТИ на основании (4.44) и (4.46) записывается как

$$K = \frac{1}{1 + \Lambda_{\text{сист}} \tau_{\text{сист}}} = \frac{1}{1 + \sum_{i=1}^n \lambda_i \tau_i}. \quad (4.48)$$

Приближенное значение КГ для высоконадежной системы, у которой $\max \lambda_i \tau_i \ll 1/n$, имеет вид

$$K \approx 1 - \sum_{i=1}^n \lambda_i \tau_i. \quad (4.49)$$

КОЭФФИЦИЕНТ ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ. Этот показатель для марковской модели строго равен произведению КГ (4.48) на ВБР (4.43). В случае высоконадежной системы приближенное выражение получается перемножением (4.44) на (4.49):

$$R(t_0) \approx \left(1 - \sum_{i=1}^n \lambda_i t_0\right) \cdot \left(1 - \sum_{i=1}^n \lambda_i \tau_i\right) \approx 1 - \sum_{i=1}^n \lambda_i (t_0 + \tau_i). \quad (4.50)$$

Замечание. Если все элементы последовательной системы с отключением при отказе имеют одинаковое время восстановления, $\mu_i = \mu$, то такая последовательная система сводится к одиночному элементу с $\lambda = \sum \lambda_i$.

4.3.3. Система без отключения при отказе. Неограниченное восстановление

Если система при отказе не отключается, то во время текущего ремонта отказавшего элемента остальные элементы системы продолжают функционировать и могут отказаться в течение времени ремонта. Это приводит к тому, что восстановление может затянуться. Простейший случай представляет собой режим неограниченного восстановления, т. е. возможность проводить независимое и одновременное восстановление любого числа отказавших элементов. Таким образом, все n элементов системы независимы полностью и в смысле отказов, и в смысле восстановления. Граф переходов подобной системы представлен на рис. 4.8.

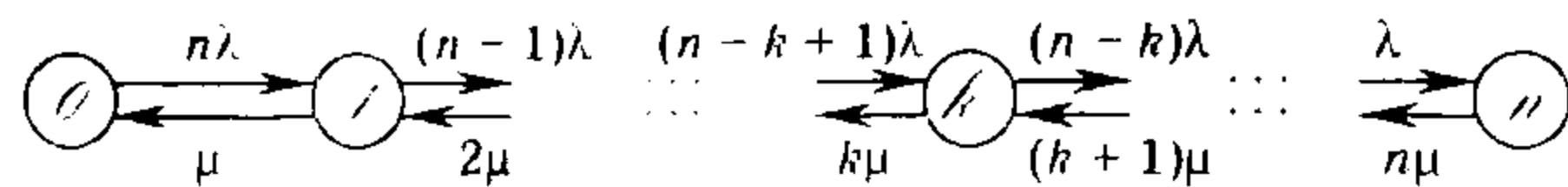


Рис. 4.8

Выражения для показателей надежности такой системы находятся легко. Многие из них совпадают с теми, которые были рассмотрены выше.

ВЕРОЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ системы та же, что и приведенная в (4.43).

СРЕДНЕЕ ВРЕМЯ БЕЗОТКАЗНОЙ РАБОТЫ. Этот показатель также совпадает с (4.45).

КОЭФФИЦИЕНТ ГОТОВНОСТИ системы в силу независимости элементов равняется произведению КГ ее элементов:

$$K = \prod_{i=1}^n K_i = \prod_{i=1}^n \frac{1}{1 + \lambda_i \tau_i}. \quad (4.51)$$

Приближенное значение КГ для высоконадежной системы, у которой $\max \lambda_i \tau_i \ll 1/n$, имеет вид

$$K \approx 1 - \sum_{i=1}^n \lambda_i \tau_i. \quad (4.52)$$

Следует заметить, что приближенные значения (4.49) и (4.52) совпадают. Это объясняется тем, что вероятностями одновременных отказов нескольких элементов высоконадежных систем можно пренебречь.

СРЕДНЕЕ ВРЕМЯ ВОССТАНОВЛЕНИЯ находится на основании формулы для КГ при известном среднем времени безотказной работы:

$$\tau_{\text{сист}} = \frac{T_{\text{сист}}(1 - K)}{K}. \quad (4.53)$$

КОЭФФИЦИЕНТ ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ Этот показатель для марковской модели равен произведению КГ (4.51) на ВБР (4.43). В случае высоконадежной системы приближенное выражение получается перемножением (4.44) на (4.52) и совпадает с (4.50):

$$R(t_0) \approx 1 - \sum_{i=1}^n \lambda_i (t_0 + \tau_i). \quad (4.54)$$

4.3.4. Система без отключения при отказе. Ограниченное восстановление

При ограниченном восстановлении отказавших элементов могут образовываться их очереди на ремонт, что приводит к увеличению времени нахождения системы в неработоспособном состоянии. Задача решается относительно легко, только если все элементы системы идентичны: здесь для строгого решения удастся использовать схему *гибели и размножения*, рассмотренную в Приложении 3. Граф переходов для случая, когда имеется k ремонтных органов, представлен на рис. 4.9.

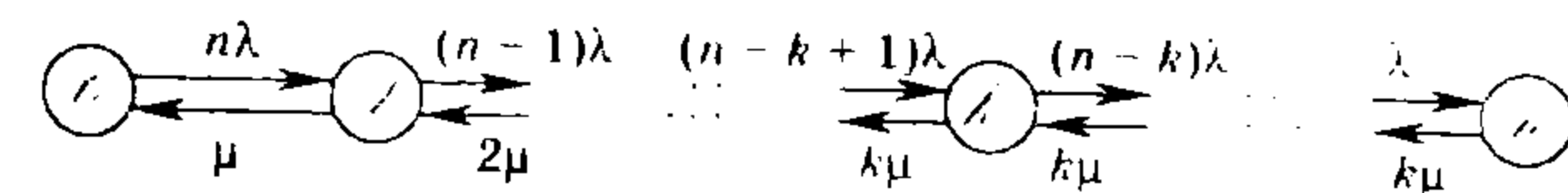


Рис. 4.9

Система алгебраических уравнений при этом имеет вид (рассматриваются условия равновесия в разрезах графа):

$$\begin{cases} n\lambda p_0 = \mu p_1, \\ (n-1)\lambda p_1 = 2\mu p_2, \\ \dots\dots\dots \\ (n-k+1)\lambda p_{k-1} = k\mu p_k, \\ (n-k)\lambda p_k = k\mu p_{k+1}, \\ \dots\dots\dots \\ \lambda p_{n-1} = k\mu p_n. \end{cases} \quad (4.55)$$

Выпишем решение, получаемое стандартным образом:

$$\begin{aligned}
 p_1 &= n\rho p_0 = \binom{n}{1}\rho p_0, \\
 p_2 &= \frac{n-1}{2}\rho p_1 = \frac{n(n-1)}{2}\rho^2 p_0 = \binom{n}{2}\rho^2 p_0, \\
 &\dots\dots\dots \\
 p_k &= \frac{n-k+1}{k}\rho p_{k-1} = \frac{n(n-1)\dots(n-k+1)}{1\cdot 2\cdot \dots\cdot k}\rho^k p_0 = \binom{n}{k}\rho^k p_0, \quad (4.56) \\
 p_{k+1} &= \frac{n-k}{k+1}\rho p_k = \frac{n(n-1)\dots(n-k)}{1\cdot 2\cdot \dots\cdot k\cdot (k+1)}\rho^{k+1} p_0 = \binom{n}{k+1}\rho^{k+1} p_0, \\
 p_n &= \binom{n}{n}\rho^n p_0 = \rho^n p_0;
 \end{aligned}$$

здесь $\rho = \lambda/\mu$.

Поскольку сумма вероятностей всех событий равна единице (условие полной вероятности), из (4.56) нетрудно найти стационарный КГ:

$$K = p_0 = \left[1 + \sum_{i=1}^n \binom{n}{i} \rho^i + \binom{n}{n} \rho^n \sum_{j=1}^{\infty} \frac{(n-k)!}{j! k^j} \rho^j \right]^{-1}. \quad (4.57)$$

Обычно в практических задачах надежности рассматриваются достаточно высоконадежные системы, у которых $\rho \ll 1/n$. В этом случае можно записать приближенное выражение:

$$K \approx 1 - n\rho. \quad (4.58)$$

Нестационарный КГ не записывается нами из-за громоздкости формулы. Кроме того, этот показатель используется при практических расчетах довольно редко.

Вероятность безотказной работы совпадает с ВБР для любой последовательной системы, в частности с (4.43).

СРЕДНЕЕ ВРЕМЯ БЕЗОТКАЗНОЙ РАБОТЫ также совпадает с аналогичным показателем для любой последовательной системы, в частности с (4.45).

СРЕДНЕЕ ВРЕМЯ ПРОСТОЯ. Эта величина может быть получена на основании известных КГ и среднего времени безотказной работы из формулы

$$\tau_{\text{сист}} = \frac{T_{\text{сист}}(1-K)}{K}. \quad (4.59)$$

§ 4.4. ДУБЛИРОВАННАЯ СИСТЕМА

Наиболее часто из резервированных систем с восстановлением встречаются дублированные системы, они будут рассмотрены с большой детальностью, даже невзирая на некоторые повторения.

Более сложные случаи — при необходимости — могут быть проанализированы с использованием общих методов, изложенных в Приложении 3.

4.4.1. Описание модели

Рассмотрим дублированную систему, состоящую из двух идентичных элементов. Обычно для марковской модели принимаются следующие допущения:

- ✓ отказы элементов наступают независимо друг от друга;
- ✓ после отказа основного элемента его функции мгновенно начинает исполнять резервный элемент;
- ✓ отказавший элемент поступает в ремонтный орган сразу же после отказа;
- ✓ восстановленный элемент по своим характеристикам неотличим от абсолютно нового элемента;
- ✓ переключение с основного элемента на резервный осуществляется мгновенно и безотказно.

При анализе надежности любой резервированной системы необходимо принять во внимание режим, в котором находится резервный элемент. Возможны три типа режима работы резервного элемента, перечислим их.

- Резервный элемент находится в тех же условиях, что и основной элемент. Такой режим называется *нагруженным* или *горячим* резервом. Интенсивности отказов соответственно резервного λ_p и основного λ элементов равны.
- Резервный элемент находится в запасе, т. е. не может отказаться. Такой режим называется *ненагруженным* или *холодным* резервом. Интенсивность отказов резервного элемента $\lambda_p = 0$.
- Резервный элемент находится в режиме, промежуточном по отношению к двум упомянутым выше. Такой режим называется *облегченным* или *теплым* резервом. Тогда $0 < \lambda_p < \lambda$. Необходимо также учесть характер восстановления отказавших элементов. Восстановление может быть ограниченным (единственный ремонтный орган) и неограниченным (в случае дублирования — два

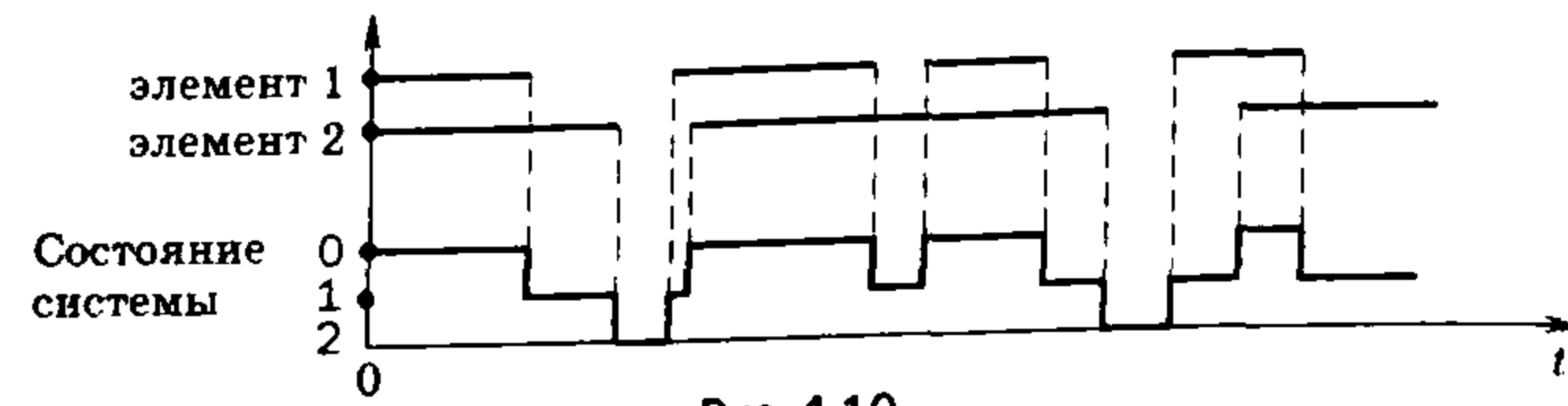


Рис. 4.10

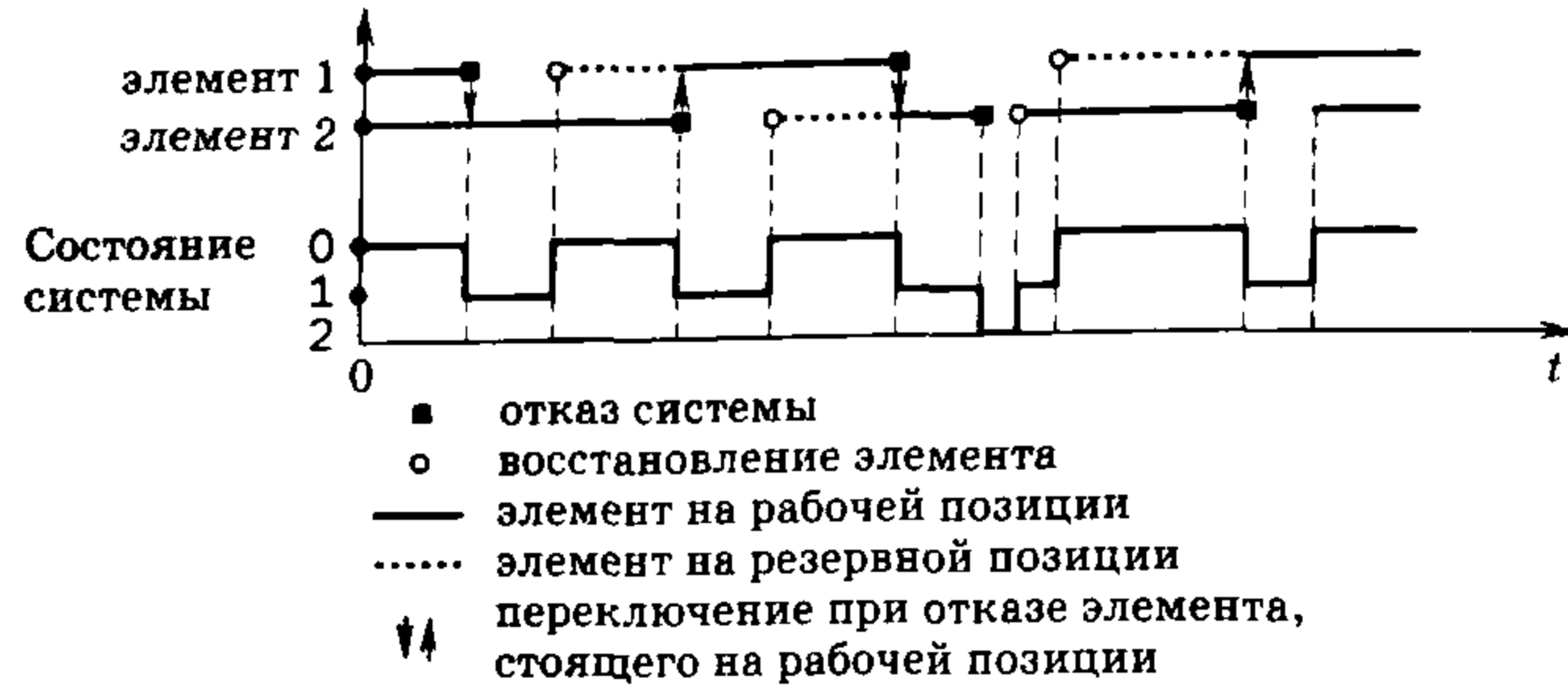


Рис. 4.11

ремонтных органа). В общем случае неограниченное восстановление подразумевает, что одновременно и независимо может восстанавливаться любое число отказавших элементов системы.

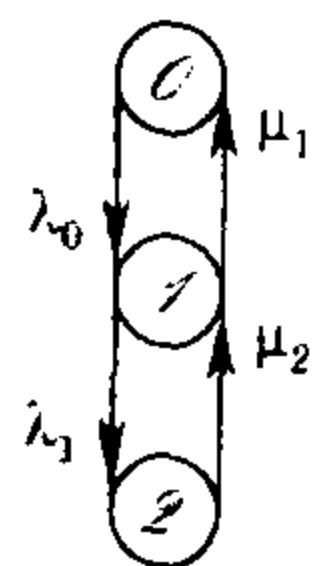


Рис. 4.12

Возможная реализация траектории поведения дублированной системы с *горячим* резервом и неограниченным восстановлением приведена на рис. 4.10, а с *холодным* — на рис. 4.11.

Граф переходов для дублированной системы в общем случае представлен на рис. 4.12.

Перейдем теперь к нахождению различных показателей надежности дублированной системы.

4.4.2. Нестационарный коэффициент готовности

Опустив детали объяснения, которые приведены в Приложении 3, запишем систему дифференциальных уравнений для вероятностей состояний графа переходов, представленного на рис. 4.11.

$$\begin{aligned} \frac{d}{dt} P_0(t) &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \\ \frac{d}{dt} P_1(t) &= \lambda_0 P_0(t) - (\lambda_1 + \mu_1) P_1(t) + \mu_2 P_2(t), \\ 1 &= P_0(t) + P_1(t) + P_2(t), \\ P_0(0) &= 1. \end{aligned} \quad (4.60)$$

Приведем систему алгебраических уравнений в преобразованиях Лапласа:

$$\begin{vmatrix} \lambda_0 + s & -\mu_1 & 0 \\ -\lambda_0 & \lambda_1 + \mu_1 + s & -\mu_2 \\ s & s & s \end{vmatrix} \times \begin{vmatrix} \varphi_0(s) \\ \varphi_1(s) \\ \varphi_2(s) \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}. \quad (4.61)$$

Поскольку в дублированной системе оба состояния 0 и 1 являются состояниями работоспособности, имеем КГ:

$$K(t) = P_0(t) + P_1(t) = 1 - P_2(t), \quad (4.62)$$

т. е. решением в преобразованиях Лапласа для $\varphi_K(s) \Leftrightarrow K(t)$ будет $\varphi_K(s) = 1/s - \varphi_2(s)$. Найдем выражение этого преобразования Лапласа:

$$\begin{aligned} \varphi_2(s) &= \frac{\begin{vmatrix} \lambda_0 + s & -\mu_1 & 1 \\ -\lambda_0 & \lambda_1 + \mu_1 + s & 0 \\ s & s & 1 \end{vmatrix}}{\begin{vmatrix} \lambda_0 + s & -\mu_1 & 1 \\ -\lambda_0 & \lambda_1 + \mu_1 + s & -\mu_2 \\ s & s & s \end{vmatrix}} = \\ &= \frac{\lambda_0 \lambda_1}{s[s^2 + s(\lambda_0 + \lambda_1 + \mu_1 + \mu_2) + \lambda_0 \lambda_1 + \lambda_0 \mu_2 + \mu_1 \mu_2]}. \end{aligned} \quad (4.63)$$

Отсюда находим искомое преобразование Лапласа

$$\begin{aligned} \varphi^*(s) &= \frac{1}{s} - \varphi_2(s) = \\ &= \frac{s^2 + s(\lambda_0 + \lambda_1 + \mu_1 + \mu_2) + \lambda_0 \mu_2 + \mu_1 \mu_2}{s[s^2 + s(\lambda_0 + \lambda_1 + \mu_1 + \mu_2) + \lambda_0 \lambda_1 + \lambda_0 \mu_2 + \mu_1 \mu_2]}. \end{aligned} \quad (4.64)$$

Дальнейшее решение идет по схеме, описанной в Приложении 3. Сначала преобразование Лапласа (4.64) представляется в виде суммы простых дробей вида

$$\varphi^*(s) = \frac{A}{s - s_1} + \frac{B}{s - s_2} + \frac{C}{s - s_3}, \quad (4.65)$$

где s_1 , s_2 и s_3 — корни знаменателя в (4.64). Затем находятся коэффициенты A , B и C для простых дробей в представлении (4.65). Заметим, что s_1 и s_2 — сопряженные корни, а $s_3 = 0$.

Опуская хотя простые, но довольно громоздкие преобразования, приходим к решению в общем виде:

$$K(t) = 1 - \frac{\lambda_0 \lambda_1}{s_1 s_2} \left[1 - \frac{1}{s_1 - s_2} \left(s_1 \exp(-s_2 t) - s_2 \exp(-s_1 t) \right) \right], \quad (4.66)$$

где $s_{1,2} = -\alpha/2 \pm \sqrt{\alpha^2/4 - \beta}$, где, в свою очередь, $\alpha = \lambda_0 + \lambda_1 + \mu_1 + \mu_2$ и $\beta = \lambda_0\lambda_1 + \lambda_0\mu_2 + \mu_1\mu_2$. Если $s_1 = s_2$, то следует применить правило Лопиталя.

Формула (4.66) позволяет получить значения $K(t)$ во всех описанных частных случаях. Для *горячего* дублирования с неограниченным восстановлением решение вообще элементарно, поскольку мы имеем два независимых элемента, а поэтому можем воспользоваться готовым решением (4.32):

$$K_{\text{сист}}(t) = 1 - (1 - K(t))^2 = 1 - \left(\frac{\lambda}{\lambda + \mu} \{1 - \exp[-(\lambda + \mu)t]\} \right)^2. \quad (4.67)$$

Отметим, что нестационарный КГ на практике применяется для анализа систем довольно редко.

4.4.3. Стационарный коэффициент готовности

При нахождении КГ можно использовать систему уравнений (4.60), положив в ней все производные равными нулю, а $P_0(t)$, $P_1(t)$ и $P_2(t)$ константами p_0 , p_1 , p_2 соответственно. Тогда можно записать систему алгебраических уравнений

$$\begin{vmatrix} -\lambda_0 & \mu_1 & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 \\ 1 & 1 & 1 \end{vmatrix} \times \begin{vmatrix} p_0 \\ p_1 \\ p_2 \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix}, \quad (4.68)$$

решением которой является

$$K = 1 - p_2 = 1 - \frac{\begin{vmatrix} -\lambda_0 & \mu_1 & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & 0 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} -\lambda_0 & \mu_1 & 1 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 \\ 1 & 1 & 1 \end{vmatrix}} = \frac{\lambda_0\lambda_1}{\lambda_0\lambda_1 + \lambda_0\mu_2 + \mu_1\mu_2},$$

или окончательно

$$K = \left(1 + \frac{\lambda_0\lambda_1}{\lambda_0\mu_2 + \mu_1\mu_2} \right)^{-1}. \quad (4.69)$$

Из (4.69) легко находятся значения КГ для различных режимов нагруженности и восстановления элементов дублированной системы. Результаты сведены в табл. 4.2.

Таблица 4.2

Резервирование	Восстановление	
	Неограниченное	Ограниченное
Нагруженное	$\frac{1}{1 + \frac{\rho^2}{1 + 2\rho}}$	$\frac{1}{1 + \frac{2\rho^2}{1 + 4\rho}}$
Ненагруженное	$\frac{1}{1 + \frac{\rho^2}{2(1 + \rho)}}$	$\frac{1}{1 + \frac{\rho^2}{1 + \rho}}$

В этом случае также, рассматривая полностью независимые элементы, когда восстановление неограниченное, а режим резервного элемента нагруженный, выражение для коэффициента готовности дублированной системы можно получить, используя непосредственно (4.33):

$$K_{\text{сист}} = 1 - (1 - K)^2 = 1 - \frac{\lambda^2}{(\lambda + \mu)^2} = \frac{2\lambda + \mu^2}{\lambda^2 + 2\lambda + \mu^2} = \frac{1}{1 + \frac{\lambda^2}{2\lambda + \mu^2}} = \frac{1}{1 + \frac{\rho^2}{1 + 2\rho}}. \quad (4.70)$$

Вариант *теплого* резерва на практике почти никогда не рассматривается из-за того, что степень облегчения режима резервного элемента не бывает сколько-нибудь точно известна. Если же необходимость оценки коэффициента готовности при этом все же возникает, то разумнее всего в таких ситуациях рассматривать показатели для *холодного* и *горячего* резерва в качестве верхней и нижней оценок соответственно.

В случае высоконадежных элементов, т. е. когда $\rho \ll 1$, на основании табл. 4.2 можно записать приближенные формулы, приведенные в табл. 4.3.

Таблица 4.3

Резервирование	Восстановление	
	Неограниченное	Ограниченное
Нагруженное	$1 - \rho^2$	$1 - 2\rho^2$
Ненагруженное	$1 - 0.5\rho^2$	$1 - \rho^2$

4.4.4. Вероятность безотказной работы

Граф переходов для расчета ВБР приведен на рис. 4.13. На этом графе состояние \mathcal{L} , соответствующее отказу системы, является поглощающим.

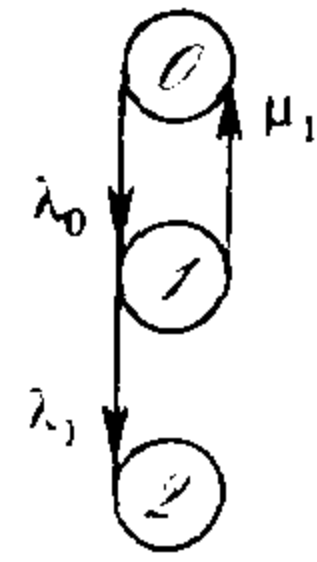


Рис. 4.13

Для этого случая имеем следующую систему дифференциальных уравнений (см. Приложение 3):

$$\begin{aligned} \frac{d}{dt} P_0(t) &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \\ \frac{d}{dt} P_1(t) &= \lambda_0 P_0(t) - (\lambda_1 + \mu_1) P_1(t), \\ P_0(0) &= 1. \end{aligned} \quad (4.71)$$

Эта система уравнений в преобразованиях Лапласа имеет вид

$$\begin{vmatrix} \lambda_0 + s & -\mu_1 \\ -\lambda_0 & \lambda_1 + \mu_1 + s \end{vmatrix} \times \begin{vmatrix} \Phi_0(s) \\ \Phi_1(s) \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \end{vmatrix}. \quad (4.72)$$

Искомое решение может быть записано в форме

$$\begin{aligned} \Phi_0(s) + \Phi_1(s) &= \frac{\begin{vmatrix} 1 & -\mu_1 \\ 0 & \lambda_1 + \mu_1 + s \end{vmatrix} + \begin{vmatrix} \lambda_0 + s & 1 \\ -\lambda_0 & 0 \end{vmatrix}}{\begin{vmatrix} \lambda_0 + s & -\mu_1 \\ -\lambda_0 & \lambda_1 + \mu_1 + s \end{vmatrix}} = \\ &= \frac{s + \lambda_0 + \lambda_1 + \mu_1}{s^2 + s(\lambda_0 + \lambda_1 + \mu_1) + \lambda_0 \lambda_1}. \end{aligned} \quad (4.73)$$

Применяя прежнюю процедуру, находим решение для ВБР при условии, что система начала работать с состояния полной работоспособности:

$$P^{(0)}(t) = \frac{1}{s_1^* - s_2^*} \left(s_1^* \exp(s_2^* t) - s_2^* \exp(s_1^* t) \right); \quad (4.74)$$

здесь $s_{1,2}^* = -\alpha^*/2 \pm \sqrt{(\alpha^*/2)^2 - \beta^*}$, $\alpha^* = \lambda_0 + \lambda_1 + \mu_1$, $\beta^* = \lambda_0 \lambda_1$, а верхний индекс в $P^{(0)}(t)$ указывает на то, что выбрано начальное условие $P_0(0) = 1$.

Далее нам пригодится значение ВБР системы при условии, что она начинает работать, будучи в состоянии \mathcal{L} , т. е. когда один элемент работоспособен, а второй восстанавливается. Это означает, что необходимо выбрать начальное условие в виде $P_1(0) = 1$.

Избегая промежуточных выкладок, запишем сразу, используя заготовку в форме (4.73):

$$\begin{aligned} \Phi_0(s) + \Phi_1(s) &= \frac{\begin{vmatrix} 0 & -\mu_1 \\ 1 & \lambda_1 + \mu_1 + s \end{vmatrix} + \begin{vmatrix} \lambda_0 + s & 0 \\ -\lambda_0 & 1 \end{vmatrix}}{\begin{vmatrix} \lambda_0 + s & -\mu_1 \\ -\lambda_0 & \lambda_1 + \mu_1 + s \end{vmatrix}} = \\ &= \frac{\lambda_0 + \mu_1}{s^2 + s(\lambda_0 + \lambda_1 + \mu_1) + \lambda_0 \lambda_1}. \end{aligned} \quad (4.75)$$

Отметим, что знаменатели в (4.73) и (4.75) одинаковы, что, впрочем, совершенно естественно, так как изменены лишь начальные условия. Обращение преобразования Лапласа (4.75) дает решение

$$P^{(1)}(t) = \frac{1}{s_1^* - s_2^*} [(s_1^* - \lambda_0 - \lambda_1) \exp(s_2^* t) - (s_2^* - \lambda_0 - \lambda_1) \exp(s_1^* t)]; \quad (4.76)$$

здесь верхний индекс $P^{(1)}(t)$ опять используется для указания на то, что выбрано начальное условие $P_1(0) = 1$.

Заметим, что в формулах отсутствует величина μ_2 , поскольку состояние \mathcal{L} является поглощающим.

4.4.5. Стационарный коэффициент интервальной готовности

Для восстанавливаемых систем, предназначенных к выполнению некоторого потока задач, данный показатель надежности представляется весьма важным. Стационарный коэффициент интервальной готовности $R(t_0)$ можно записать с использованием формулы полной вероятности в виде

$$R(t_0) = p_0 P^{(0)}(t) + p_1 P^{(1)}(t), \quad (4.77)$$

где p_0 и p_1 — стационарные вероятности состояний \mathcal{O} и \mathcal{L} , а $P^{(0)}(t)$ и $P^{(1)}(t)$ определяются по формулам (4.74) и (4.76) соответственно. Стационарные вероятности p_0 и p_1 находятся на базе системы уравнений (4.71):

$$p_0 = 1 - \frac{\begin{vmatrix} 0 & \mu_1 & 0 \\ 0 & -(\lambda_1 + \mu_1) & \mu_2 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} -\lambda_0 & \mu_1 & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 \\ 1 & 1 & 1 \end{vmatrix}} = \frac{\mu_1 \mu_2}{\lambda_0 \lambda_1 + \lambda_0 \mu_2 + \mu_1 \mu_2}, \quad (4.78)$$

$$P_1 = 1 - \frac{\begin{vmatrix} -\lambda_0 & \mu_1 & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} -\lambda_0 & 0 & 0 \\ \lambda_0 & 0 & \mu_2 \\ 1 & 1 & 1 \end{vmatrix}} = \frac{\lambda_0 \mu_2}{\lambda_0 \lambda_1 + \lambda_0 \mu_2 + \mu_1 \mu_2}. \quad (4.79)$$

4.4.6. Среднее время безотказной работы

Среднее время работы до отказа может быть найдено из преобразования Лапласа (4.73):

$$T^{(0)} = \frac{s + \lambda_0 + \lambda_1 + \mu_1}{s^2 + s(\lambda_0 + \lambda_1 + \mu_1) + \lambda_0 \lambda_1} \Big|_{s=0},$$

или окончательно

$$T^{(0)} = \frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_0 \lambda_1}. \quad (4.80)$$

Среднее время работы между отказами получаем из преобразования Лапласа (4.75):

$$T^{(1)} = \frac{\lambda_0 + \mu_1}{s^2 + s(\lambda_0 + \lambda_1 + \mu_1) + \lambda_0 \lambda_1} \Big|_{s=0},$$

или окончательно

$$T^{(1)} = \frac{1}{\lambda_0} + \frac{\mu_1}{\lambda_0 \lambda_1}. \quad (4.81)$$

Обратим внимание на различие в (4.80) и (4.81): среднее время работы до отказа $T^{(0)}$ больше, чем среднее время работы между отказами $T^{(1)}$, на величину $1/\lambda_1$. Действительно, система из состояния 0 переходит в состояние 1 в среднем через время $1/\lambda_1$, а затем уже продолжает свой процесс переходов из состояния в состояние, пока не попадет в поглощающее состояние 2.

Можно получить тот же результат и с помощью иного подхода, используя рекуррентные формулы. Из только что приведенных рассуждений следует, что справедливо следующее равенство:

$$T^{(0)} = \frac{1}{\lambda_0} + T^{(1)}. \quad (4.82)$$

Перейдя в состояние 1, система пребывает в нем в среднем время $1/(\lambda_1 + \mu_1)$, после чего она либо с вероятностью $\lambda_1/(\lambda_1 + \mu_1)$ перехо-

дит в поглощающее состояние 2, либо с вероятностью $\mu_1/(\lambda_1 + \mu_1)$ возвращается в состояние 0, т. е. можно записать рекуррентное соотношение

$$T^{(1)} = \frac{1}{\lambda_1 + \mu_1} + \frac{\mu_1}{\lambda_1 + \mu_1} T^{(0)}. \quad (4.83)$$

Подставив (4.83) в (4.82), получаем выражение для среднего времени работы до отказа:

$$T^{(0)} = \frac{1}{1 - \frac{1}{\lambda_1 + \mu_1}} \left(\frac{1}{\lambda_0} + \frac{1}{\lambda_1 + \mu_1} \right) = \frac{\lambda_0 + \lambda_1 + \mu_1}{\lambda_0 \lambda_1} = \frac{1}{\lambda_0} + \frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_0 \lambda_1}. \quad (4.84)$$

Из (4.84) и (4.82) следует, что среднее время между отказами равно

$$T^{(1)} = T^{(0)} - \frac{1}{\lambda_0} = \frac{1}{\lambda_1} + \frac{\mu_1}{\lambda_0 \lambda_1}. \quad (4.85)$$

Заметим, что здесь, как и при расчете ВБР, тип восстановления не играет роли, поскольку имеется всего два состояния 0 и 1. При расчете таких показателей надежности, как ВБР и среднее время безотказной работы, должно учитываться число ремонтных органов, поскольку от общего числа ремонтных органов зависит значение времени простоя системы в неработоспособном состоянии.

Окончательные результаты четырех частных случаев в дублированной системе сведены в табл. 4.4. В ней для ненагруженного и нагруженного резерва соответственно имеем

$$\lambda_0 = \lambda_1 = \lambda,$$

$$\lambda_0 = 2\lambda, \quad \lambda_1 = \lambda.$$

Если рассматриваются высоконадежные элементы, первым слагаемым во всех формулах табл. 4.4 можно пренебречь, так как $\lambda \ll \mu$. Более того, различие между $T^{(0)}$ и $T^{(1)}$ в пределе пропадает.

Таблица 4.4

Резервирование	$T^{(0)}$ (до отказа)	$T^{(1)}$ (между отказами)
Нагруженное	$\frac{3}{2\lambda} + \frac{\mu}{2\lambda^2}$	$\frac{1}{\lambda} + \frac{\mu}{2\lambda^2}$
Ненагруженное	$\frac{2}{\lambda} + \frac{\mu}{\lambda^2}$	$\frac{1}{\lambda} + \frac{\mu}{\lambda^2}$

Таблица 4.5

Резервирование	$T^{(0)}$ или $T^{(1)}$
Нагруженное	$\frac{\mu}{2\lambda^2}$
Ненагруженное	$\frac{\mu}{\lambda^2}$

В табл. 4.5 представлены приближенные формулы для средних времен до отказа и между отказами в высоконадежной дублированной системе.

§ 4.5. МЕТОД ВВЕДЕНИЯ ФИКТИВНЫХ СОСТОЯНИЙ

Описанные методы оценки показателей надежности резервированных систем с восстановлением пригодны для систем, у которых время пребывания в каждом состоянии имеет экспоненциальное распределение. Однако это допущение далеко не всегда оказывается приемлемым. Чтобы в рамках марковской модели описать распределения с возрастающей или убывающей интенсивностью отказов, используется искусственный прием: вводятся дополнительные фиктивные состояния.

При помощи таких фиктивных состояний удастся ввести в модель системы эрланговское (из класса *стареющих*) и гиперэкспоненциальное (из класса *молодеющих*) распределения. Например, если время восстановления элемента может быть аппроксимировано распределением Эрланга второго порядка общего вида, то марковский граф переходов может быть представлен схемой рис. 4.14.

В этом графе переходов время восстановления представляет собой сумму двух случайных величин с экспоненциальными распределениями, одно из них с параметром μ , а другое — с параметром μ' . Распределение времени безотказной работы экспоненциальное с параметром λ .

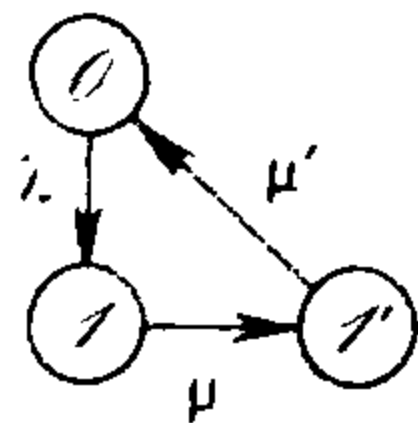


Рис. 4.14

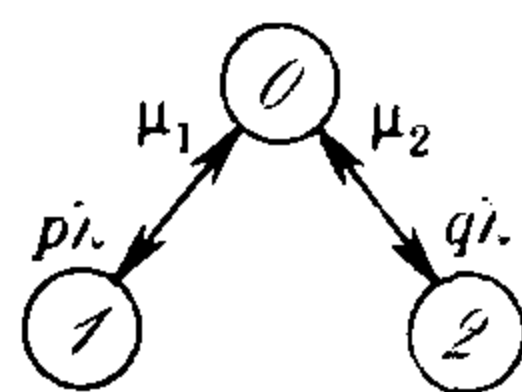


Рис. 4.15

На рис. 4.15 представлен граф переходов для элемента, у которого время восстановления является простейшим гиперэкспоненциальным, представляющим собой взвешенную сумму двух экспонент. Ради компактности на рисунке пе-

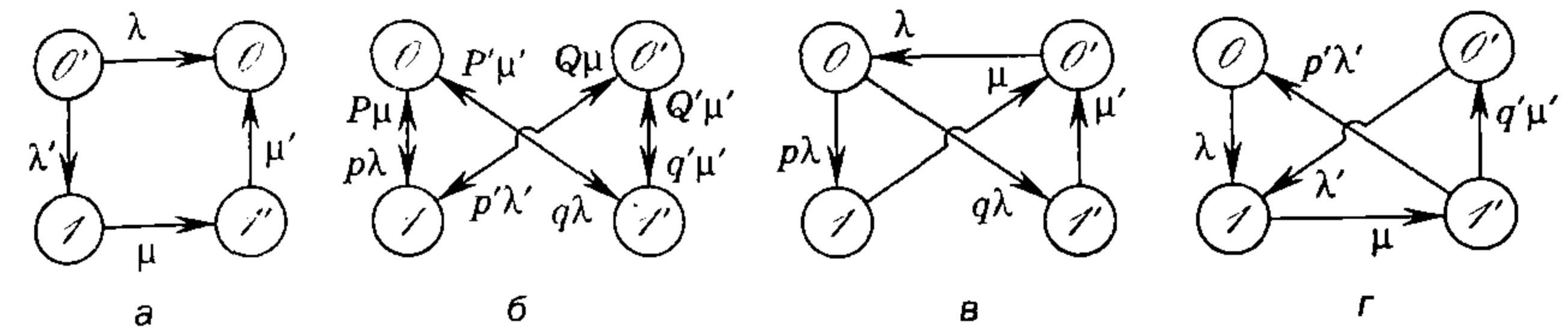


Рис. 4.16

реходы обозначены двунаправленными стрелками. На рис. 4.16 приведены другие варианты графов переходов для следующих случаев:

- (а) эрланговское распределение времени безотказной работы и эрланговское распределение времени восстановления;
- (б) гиперэкспоненциальное распределение времени безотказной работы и гиперэкспоненциальное распределение времени восстановления;
- (в) эрланговское распределение времени безотказной работы и гиперэкспоненциальное распределение времени восстановления;
- (г) гиперэкспоненциальное распределение времени безотказной работы и эрланговское распределение времени восстановления.

Из этих примеров следует, что введение фиктивных состояний позволяет ввести аппроксимирующие распределения для случаев возрастающей и убывающей интенсивности отказов, но, упрощая задачу с одной стороны (немарковская модель аппроксимируется марковской), мы усложняем ее, увеличивая общее число состояний.

Принцип нахождения показателей надежности остается прежним. Заметим, что использование фиктивных состояний даже в случае дублирования приводит к громоздким результатам, поэтому мы опускаем их детальное рассмотрение.

Глава 5

Стареющие элементы и системы

На практике часто не известен вид распределения времени до отказа, но из опыта известна информация о среднем значении. Если предположить, что элемент или система обладают свойством *старения*, то на основании этой весьма ограниченной информации можно получить граничные оценки для показателей надежности.

Данная глава является факультативной.

§ 5.1. МОНОТОННАЯ ФУНКЦИЯ ИНТЕНСИВНОСТИ ОТКАЗОВ

Большая часть технических объектов обладает естественным свойством: с течением времени система или элементы, ее составляющие, стареют, т. е. их характеристики надежности ухудшаются. Это означает, что уже проработавший какое-то время технический объект отличается от нового в худшую сторону. Конечно, может наблюдаться противоположный феномен: если технический объект уже проработал период возможной «детской смертности», то он оказывается более предпочтительным, чем новый: в нем может произойти *выжигание* слабых элементов или их *приработка* и т. п. Однако ниже мы будем рассматривать только *стареющие* системы как наиболее важные для инженерной практики.

Само старение может характеризоваться ухудшением различных свойств надежности. Так, с физической точки зрения процесс старения наиболее естественно характеризуется возрастанием функции интенсивности отказов $\lambda(t)$ (рис. 5.1).

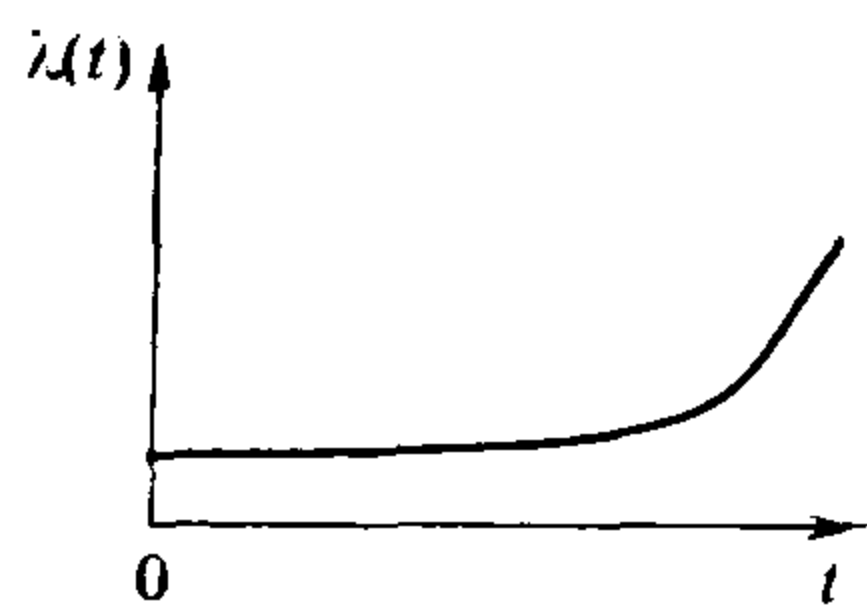


Рис. 5.1

Будем называть для краткости такие распределения *стареющими* распределениями, а элементы или системы, обладающие таким типом старения, соответственно *стареющими* элементами и *стареющими* системами.

Есть и другие характеристики старения, например, распределение может обладать возрастающей в среднем функцией интенсивности отказов. Или же среднее остаточное время безотказной работы (ВБР) с течением времени убывает. Но не будем останавливаться на столь экзотических видах распределений.

Если нам вообще не известен вид функции распределения до отказа, то, пользуясь известным из теории вероятностей неравенством Чебышева, мы можем для любого распределения задать следующую верхнюю границу:

$$\Pr\{|X - E\{X\}| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}, \quad (5.1)$$

где X — случайная величина, $E\{X\}$ и σ^2 — ее математическое ожидание и дисперсия, а ε — произвольная положительная константа.

Приведем известное доказательство этого результата. По определению

$$\Pr\{|X - E\{X\}| \geq \varepsilon\} = \int_{|x - E\{X\}| \geq \varepsilon} dF(x). \quad (5.2)$$

Поскольку область интегрирования $(1/\varepsilon)|x - E\{X\}| \geq 1$, можно записать:

$$\begin{aligned} \int_{|x - E\{X\}| \geq \varepsilon} dF(x) &\leq \frac{1}{\varepsilon^2} \int_{|x - E\{X\}| \geq \varepsilon} (x - E\{X\})^2 dF(x) \leq \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x - E\{X\})^2 dF(x) = \frac{\sigma^2}{\varepsilon^2}, \end{aligned} \quad (5.3)$$

что и требовалось доказать.

Из формулы (5.3) видно, что при ее универсальности оценка, получаемая с ее помощью, дает весьма грубое приближение и к тому же имеет смысл только для значений аргумента, лежащих на расстоянии более σ от среднего. Оказывается, могут быть получены улучшенные оценки для случая, когда имеется некоторая дополнительная информация о распределении времени до отказа.

§ 5.2. СТАРЕЮЩИЙ ЭЛЕМЕНТ

Если мы знаем, что функция интенсивности отказов возрастает во времени, то это позволяет построить верхнюю и нижнюю границы для ВБР, если известно значение среднего времени работы до отказа.

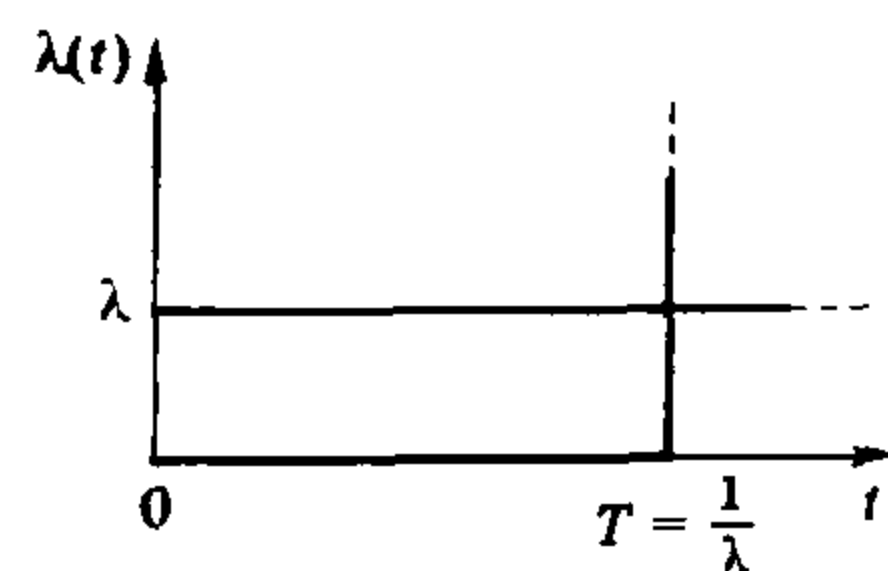


Рис. 5.2

При дальнейших выводах мы будем часто опираться на то, что экспоненциальное и вырожденное* распределения являются граничными в классе стареющих распределений. Действительно, первое из них имеет постоянную интенсивность отказов, а второе — дельта-функцию. На рис. 5.2 изображены интенсивности

отказов для случая, когда у обоих распределений одинаковое среднее, равное T .

5.2.1. Границы для вероятности безотказной работы

нижняя граница. Рассмотрим три распределения с одинаковыми средними, равными T : *стареющее*, вырожденное и экспоненциальное с параметром $1/T$. Обозначим для этих трех распределений функции ВБР соответственно через $P(t)$, $G(t)$ и $E(t)$. По рис. 5.3 видно, что $G(t)$ пересекается с $E(t)$ и $P(t)$ в точке с абсциссой T . Можно показать, что точка пересечения функций $E(t)$ и $P(t)$ лежит правее точки T . Из условия равенства средних следует, что площади, ограниченные каждой из кривых $G(t)$, $E(t)$ и $P(t)$ и осью абсцисс равны. Из этого можно сделать вывод, что для точки t^* — абсциссы той, в которой пересекаются функции $E(t)$ и $P(t)$, выполняется условие

$$\int_0^{t^*} [P(t) - E(t)] dt = \int_{t^*}^{\infty} [E(t) - P(t)] dt. \quad (5.4)$$

Из рассмотрения рис. 5.3 видно, что, поскольку $G(t) > P(t)$ при $t < T$,

$$\int_0^T [P(t) - E(t)] dt < \int_0^T [G(t) - E(t)] dt = \int_{t^*}^{\infty} [E(t) - P(t)] dt, \quad (5.5)$$

откуда с очевидностью следует, что должно выполняться условие $t^* > T$.

Таким образом, нижняя граница ВБР для любого *стареющего* распределения со средним T определяется как

$$P(t) \geq \begin{cases} \exp(-t/T) & \text{для } t < T, \\ 0 & \text{для } t \geq T. \end{cases} \quad (5.6)$$

* Вырожденное распределение, соответствующее постоянной величине (см. Приложение 2).

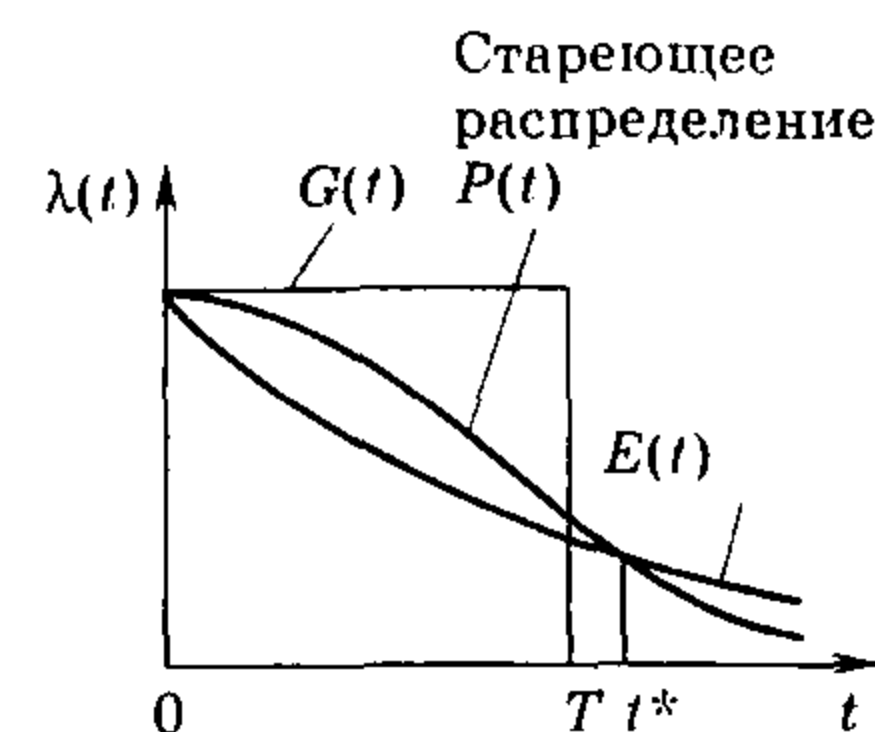


Рис. 5.3



Рис. 5.4

ВЕРХНЯЯ ГРАНИЦА. При $t \leq T$ определение верхней границы для ВБР тривиально: $P(t) \leq 1$. Нахождение верхней границы при $t > T$ требует вспомогательных построений. Рассмотрим семейство усеченных экспоненциальных функций $E_k^*(t)$ таких, что у всех них средние равны T . *Стареющая* функция $P(t)$ пересекает любую из экспоненциальных функций $E_k^*(t)$, усеченную справа в точке t_k дважды: сначала сверху вниз левее точки t_k , а затем в точке t_k (рис. 5.4). Это следует из того, что обе функции имеют одно и то же среднее, т. е. равны площади

$$\int_{t: P(t) > E_k^*(t)} [P(t) - E_k^*(t)] dt = \int_{t: E_k^*(t) > P(t)} [E_k^*(t) - P(t)] dt. \quad (5.7)$$

На рис. 5.4 показаны исходная стареющая функция $P(t)$, вырожденная функция распределения $G(t)$, семейство экспоненциальных функций $E_1(t)$, $E_2(t)$, ... и соответствующие им усеченные справа экспоненциальные функции $E_1^*(t)$, $E_2^*(t)$, ... Усеченные экспоненциальные функции таковы, что их средние равны среднему для распределения $P(t)$, т. е. T . Из построения видно, что любая усеченная в точке t^* экспоненциальная функция лежит выше функции $P(t)$ в той же точке, т. е. $E^*(t^*) < P(t^*)$. Осталось найти для каждой усеченной экспоненты соответствующие показатели. Выберем произвольную экспоненциальную функцию, пересекающую функцию $P(t)$ снизу, например $E_2(t)$ на рис. 5.4. Найдем соответствующую ей усеченную функцию $E_2^*(t)$, которая характеризуется показателем экспоненты ω_2 и точкой усечения t_2^* . Как уже отмечалось, должно выполняться условие

$$\int_0^{t_2^*} E_2^*(t) dt = \int_0^{\infty} P(t) dt = T. \quad (5.8)$$

Это же условие можно записать в виде

$$\int_0^{t_2^*} \exp(-\omega_2 t) dt = T, \quad (5.9)$$

где параметр экспоненты ω_2 связан с точкой усечения t_2^* уравнением, эквивалентным (5.9):

$$1 - \omega_2 T = \exp(-\omega_2 t_2^*). \quad (5.10)$$

Таким образом может быть построена вся непрерывная функция, соответствующая верхней границе для ВБР:

$$P(t) \leq \begin{cases} 1 & \text{для } t \leq T, \\ \exp(-\omega_k t) & \text{для } t > T, \end{cases} \quad (5.11)$$

где ω_k зависит от t_k и находится из уравнения типа (5.10).

Заметим, что для практических расчетов нас обычно больше интересует гарантированный результат, т. е. нижняя граница возможного значения ВБР.

5.2.2. Отношения между стареющими и экспоненциальными распределениями

Рассмотрим функцию

$$\Lambda(t) = \int_0^t \lambda(x) dx. \quad (5.12)$$

Для экспоненциального распределения эта функция равна

$$\Lambda_{\text{эксп}}(t) = \lambda t. \quad (5.13)$$

Мы уже отмечали, что при любом законе распределения времени работы до отказа справедливо выражение

$$P(t) = e^{-\Lambda(t)}. \quad (5.14)$$

Из (5.14) следует, что $\Lambda(t) = -\ln P(t)$. Это означает, что если мы хотим изучить взаимное поведение стареющей и экспоненциальных функций, то достаточно рассмотреть таковое на функциях $\Lambda(t)$.

Возможны следующие типы поведения функции $\Lambda(t)$ для стареющих распределений (рис. 5.5):

- (1) $\Lambda(t) < \lambda t$ для всех t — стареющая функция $P(t)$ всегда лежит выше экспоненциальной функции $E(t)$;
- (2) $\Lambda(t) > \lambda t$ для всех t — $\Lambda(t)$ выпукла, $E(t)$ всегда лежит выше стареющей функции $P(t)$;

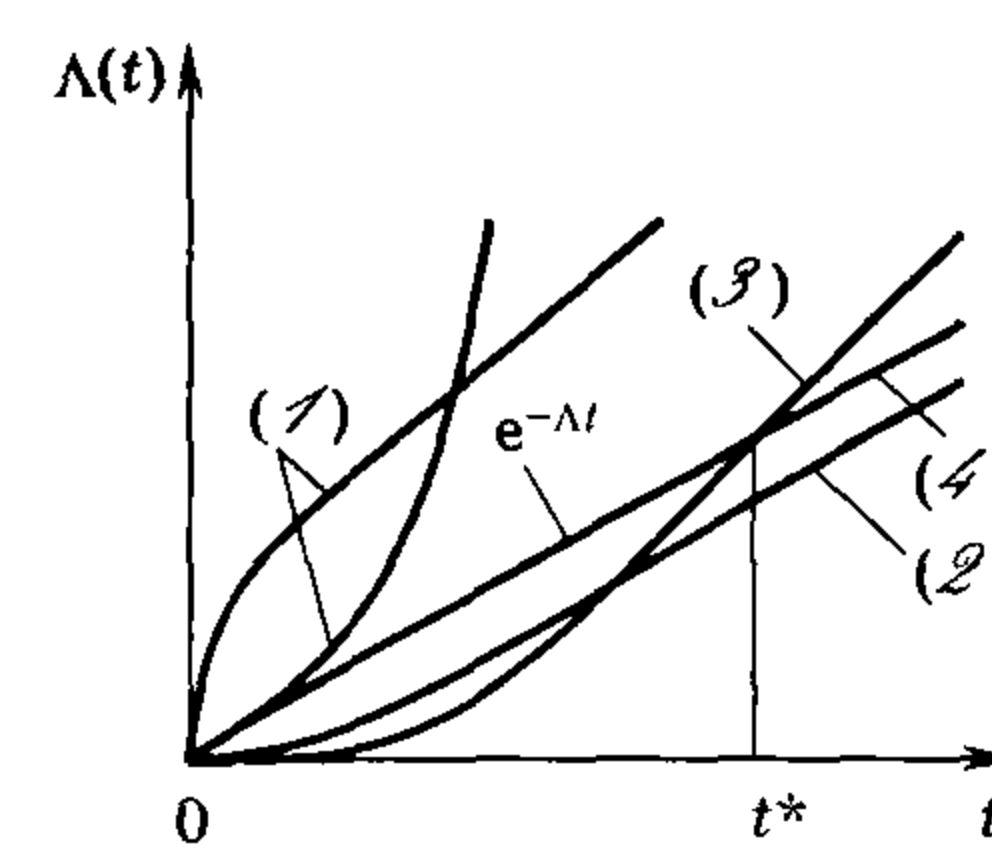


Рис. 5.5

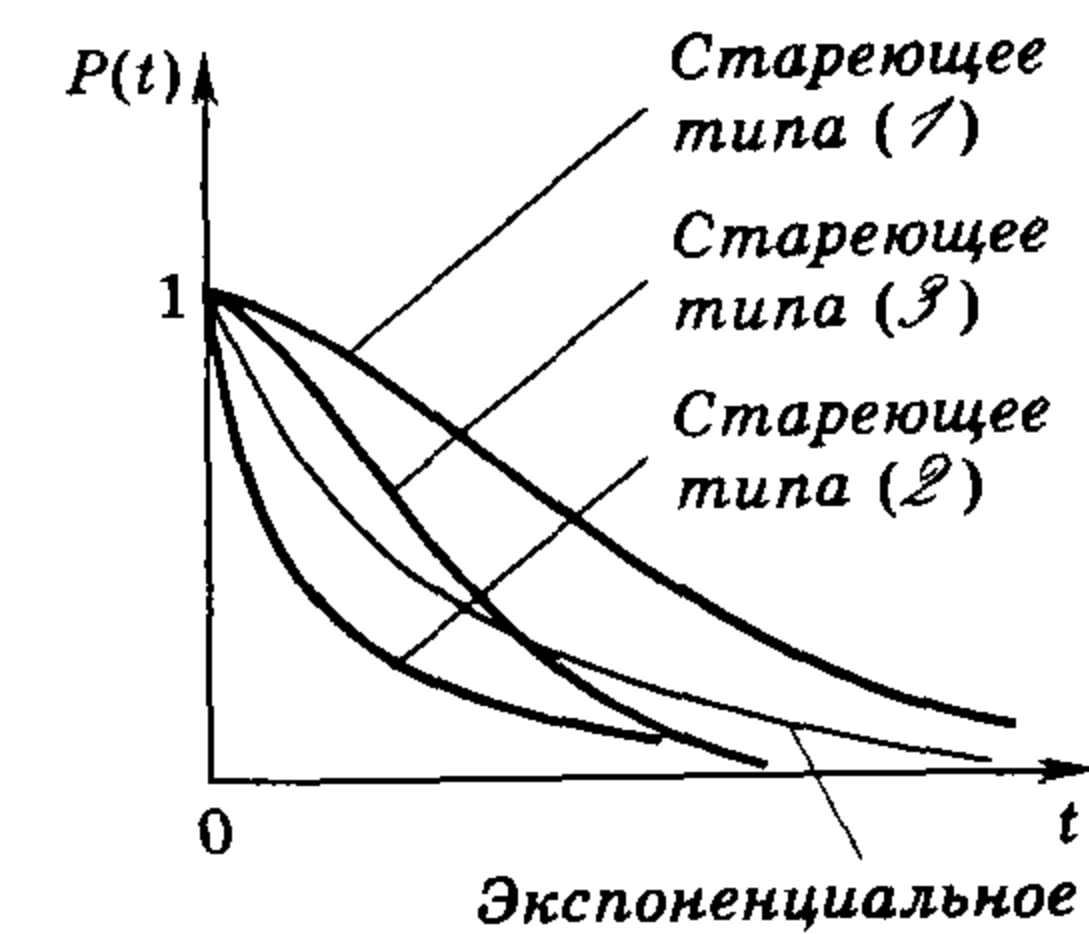


Рис. 5.6

(3) $\Lambda(t) > \lambda t$ для всех t — $\Lambda(t)$ вогнута, $E(t)$ всегда лежит выше стареющей функции $P(t)$;

(4) $\Lambda(t) < \lambda t$ для всех $t < t^*$ и $\Lambda(t) > \lambda t$ для всех $t > t^*$ — $P(t)$ пересекает экспоненциальную функцию сверху вниз в точке t^* , причем такое пересечение возможно лишь однажды.

Избегая скучных строгих рассуждений, мы здесь фактически прибегли к древнегреческому способу доказательства: «смотри чертеж».

Итак, стареющая функция $P(t)$ пересекает экспоненциальную функцию не более одного раза, причем если такое пересечение происходит в момент t^* , то происходит сверху вниз, т. е.

$$P(t) = \begin{cases} \geq \exp(-\lambda t) & \text{для } t \leq t^*, \\ \leq \exp(-\lambda t) & \text{для } t \geq t^*. \end{cases} \quad (5.15)$$

Иначе говоря, может быть только три вида отношений между стареющим и экспоненциальным распределениями, которые представлены на рис. 5.6.

Для всякого стареющего распределения функция $\Lambda(t) = -\ln P(t)$ выпукла, т. е. можно записать:

$$\frac{\ln P(t) - \ln P(0)}{t - 0} = \frac{1}{t} \ln \frac{P(t)}{P(0)} \downarrow \text{ при } t \rightarrow \infty. \quad (5.17)$$

Смысл знака \downarrow заключается в том, что предшествующее ему выражение уменьшается.

Подставив $P(0) = 1$, получаем следующее утверждение:

$$\sqrt[t]{P(t)} \downarrow \text{ при } t \rightarrow \infty. \quad (5.18)$$

Это означает, например, что если для стареющей системы мы знаем некоторое значение $P(t)$, то при всяком $\gamma > 1$ имеет место $P(\gamma t) < [P(t)]^{1/\gamma}$. Тогда при всех $x > t$ можно записать неравенство

$$P(x) \leq [P(t)]^{x/t}. \quad (5.19)$$

На практике неравенство (5.19) играет важную роль: если мы знаем значение вероятности безотказной работы для некоторого **малого** t , то легко найти гарантированную верхнюю границу для большего интервала времени.

5.2.3. Граница для вероятности безотказной работы по результатам испытаний на начальном участке

В ряде случаев относительно большая партия элементов может быть испытана в течение короткого времени. Вычисленное значение интенсивности отказов принимается как истинное. Насколько справедливо такое предположение, если на основании здравого смысла или наблюдений над аналогами известно, что распределение времени работы до отказа является *стареющим*?

Допустим, что в результате таких испытаний получено значение интенсивности отказов λ , которое можно считать начальным,

$$\lambda(0) = \frac{n}{N\Delta}, \quad (5.20)$$

где n — число отказов за время Δ , N — число испытываемых элементов, Δ — длительность испытаний.

Предположение об экспоненциальности распределения с интенсивностью отказов (5.20) приводит к тому, что мы будем завышать истинное значение вероятности безотказной работы, поскольку у рассматриваемого *стареющего* распределения производная при $t = 0$ совпадает с производной экспоненты, а при $t > 0$ функция $P(t)$ лежит всюду ниже экспоненциальной функции (см. случаи \mathcal{L} и \mathcal{J} на рис. 5.6).

5.2.4. Конечность моментов

Начальные моменты всех порядков всякого *стареющего* распределения конечны. Действительно, на основании (5.19) для всякого t можно записать:

$$\int_0^\infty x^r P(x) dx \leq \int_0^\infty x^r [P(t)]^{x/t} dx = \int_0^\infty x^r e^{-\beta x} dx < \int_0^\infty x^r e^{-\beta x} dx < \infty, \quad (5.21)$$

где $(P(t))^{1/t}$ обозначено через $e^{-\beta}$.

Из теории вероятностей известно, что экспоненциальное распределение имеет конечные моменты всех порядков, а (5.21) показывает, что моменты любого порядка для *стареющего* распределения всегда меньше соответствующих моментов для экспоненциального распределения. Это означает, в частности, что *стареющие* распределения имеют коэффициент вариации меньший, чем экспоненциальное.

5.2.5. Границы, выраженные через квантиль

Пусть γ_p есть квантиль* распределения $P(t)$, т. е. $\Pr\{x > \gamma_p\} = 1 - p$. Всегда можно найти экспоненциальную функцию, проходящую через заданную точку $(\gamma_p, 1 - p)$. Такая экспонента будет иметь вид (рис. 5.7)

$$\exp(-\alpha\gamma_p) = 1 - p. \quad (5.22)$$

Тогда

$$P(t) = \begin{cases} \geq \exp(-\alpha\gamma_p) & \text{для } t \leq \gamma_p, \\ \leq \exp(-\alpha\gamma_p) & \text{для } t \geq \gamma_p, \end{cases} \quad (5.23)$$

где

$$\alpha = -\frac{\ln(1-p)}{\gamma_p}. \quad (5.24)$$

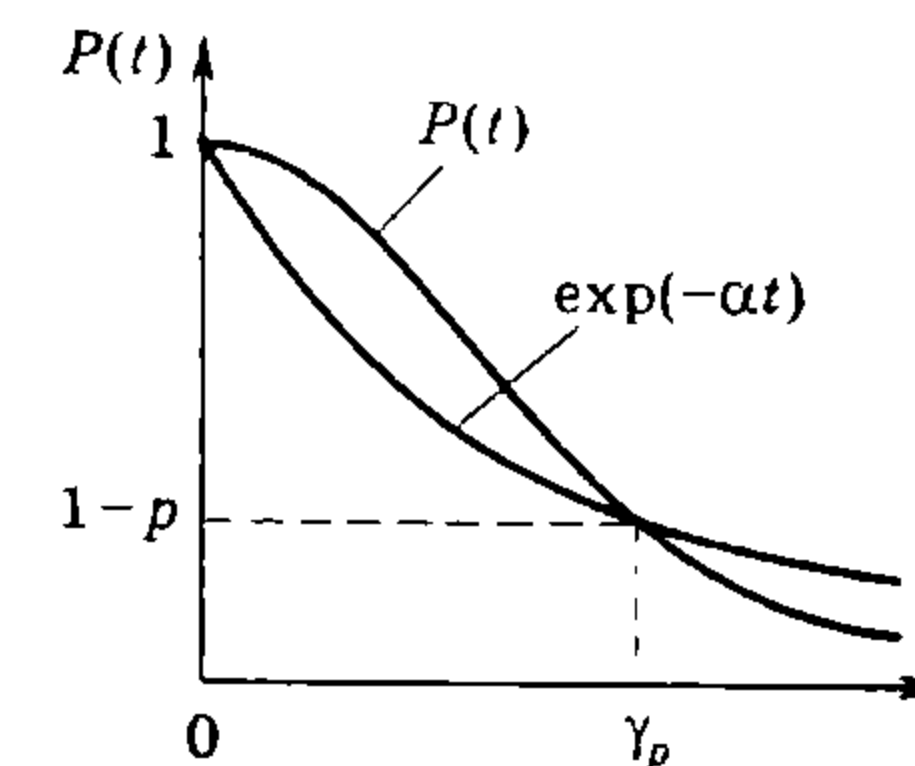


Рис. 5.7

Этот результат важен при практических расчетах: если известно значение ВБР в некоторый момент времени t и справедливо предположение о *старении* элемента, то внутри интервала $[0, t]$ экспонента с показателем α , определенным в (5.24), будет гарантированной нижней оценкой для вероятности $P(t)$.

§ 5.3. ПОСЛЕДОВАТЕЛЬНАЯ СИСТЕМА ИЗ СТАРЕЮЩИХ ЭЛЕМЕНТОВ

5.3.1. Предварительная лемма

Предварительно докажем лемму, которая окажется необходимой в последующем изложении.

* Квантиль — одна из числовых характеристик случайных величин, применяемых в математической статистике. Она представляет собой значение аргумента функции распределения для некоторого фиксированного уровня вероятности.

Пусть для функций $f(x)$ и $g(x)$ выполняются условия:

- ✓ функция $f(x)$ монотонная ограниченная и неотрицательная на всей положительной полуоси;
- ✓ функция $g(x)$ абсолютно интегрируема на всей положительной полуоси;
- ✓ функция $g(x)$ такова, что $g(x) \geq 0$ для $x < a$ и $g(x) < 0$ для $x > a$ и дополнительно функция $g(x)$ удовлетворяет условию

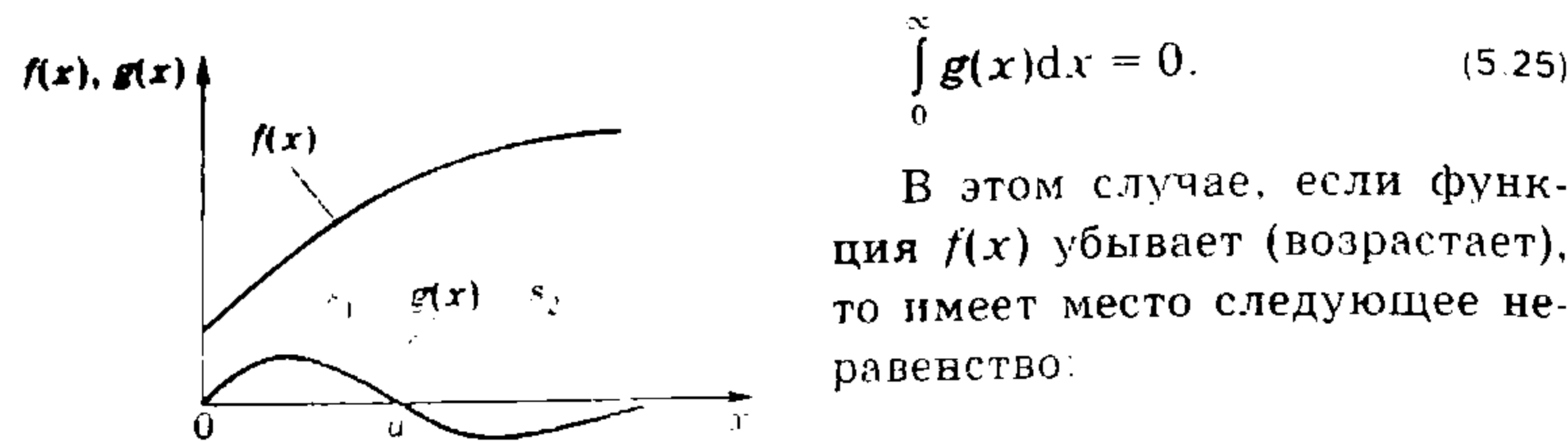


Рис 5.8

$$\int_0^{\infty} g(x) dx = 0. \quad (5.25)$$

В этом случае, если функция $f(x)$ убывает (возрастает), то имеет место следующее неравенство:

$$\int_0^{\infty} f(x)g(x)dx \leq (\geq) 0. \quad (5.26)$$

Запишем цепочку простых преобразований:

$$\begin{aligned} \int_0^{\infty} f(x)g(x)dx &= \int_0^a f(x)g(x)dx + \int_a^{\infty} f(x)g(x)dx \leq (\geq) \int_0^a [\max_{0 \leq x \leq a} f(x)]g(x)dx + \\ &+ \int_a^{\infty} [\min_{0 \leq x \leq a} f(x)]g(x)dx = f(a) \int_0^a g(x)dx + f(a) \int_a^{\infty} g(x)dx = f(a) \int_0^{\infty} g(x)dx = 0. \end{aligned} \quad (5.27)$$

Смысл этой теоремы легко понять из рассмотрения рис. 5.8, где представлен случай возрастающей функции $f(x)$. В исходном интеграле площадь s_1 берется с меньшим весом, чем площадь s_2 .

Вероятность безотказной работы

Рассмотрим последовательную систему из n элементов. Перенумеруем эти элементы в порядке возрастания значений среднего времени до отказа: $T_1 \leq T_2 \leq \dots \leq T_n$. В последовательных системах из *стареющих* элементов могут быть получены границы ВБР, базирующиеся на аналогичных границах для *стареющих* элементов.

нижняя граница Нижняя оценка ВБР для каждого элемента, как было показано, определяется по (5.6). Тогда ниж-

няя граница ВБР последовательной системы $\underline{P}_{\text{сист}}(t)$ определится как

$$\underline{P}_{\text{сист}}(t) = \begin{cases} E(t) = \prod_{i=1}^n E_i(t) & \text{для } t \leq \min_{1 \leq i \leq n} T_i, \\ 0 & \text{для } t > \min_{1 \leq i \leq n} T_i, \end{cases} \quad (5.28)$$

где $E_i(t) = \exp(-t/T_i)$, а T_i — среднее время безотказной работы i -го элемента. На рис. 5.9 приведено пояснение для случая $n = 3$.

Воспользовавшись (5.6), можно записать выражение нижней границы ВБР последовательной системы из *стареющих* элементов как

$$\underline{P}_{\text{сист}}(t) \geq \begin{cases} \exp(-t \sum (T_i)^{-1}) & \text{для } t < T_{\min}, \\ 0 & \text{для } t \geq T_{\min}. \end{cases} \quad (5.29)$$

ВЕРХНЯЯ ГРАНИЦА. Используя верхнюю границу ВБР *стареющего* элемента (5.11), можно построить верхнюю границу ВБР последовательной системы, состоящей из n независимых *стареющих* элементов. Избегая общих рассуждений, продемонстрируем метод получения этой границы с помощью последовательной системы, состоящей из трех элементов (рис. 5.10).

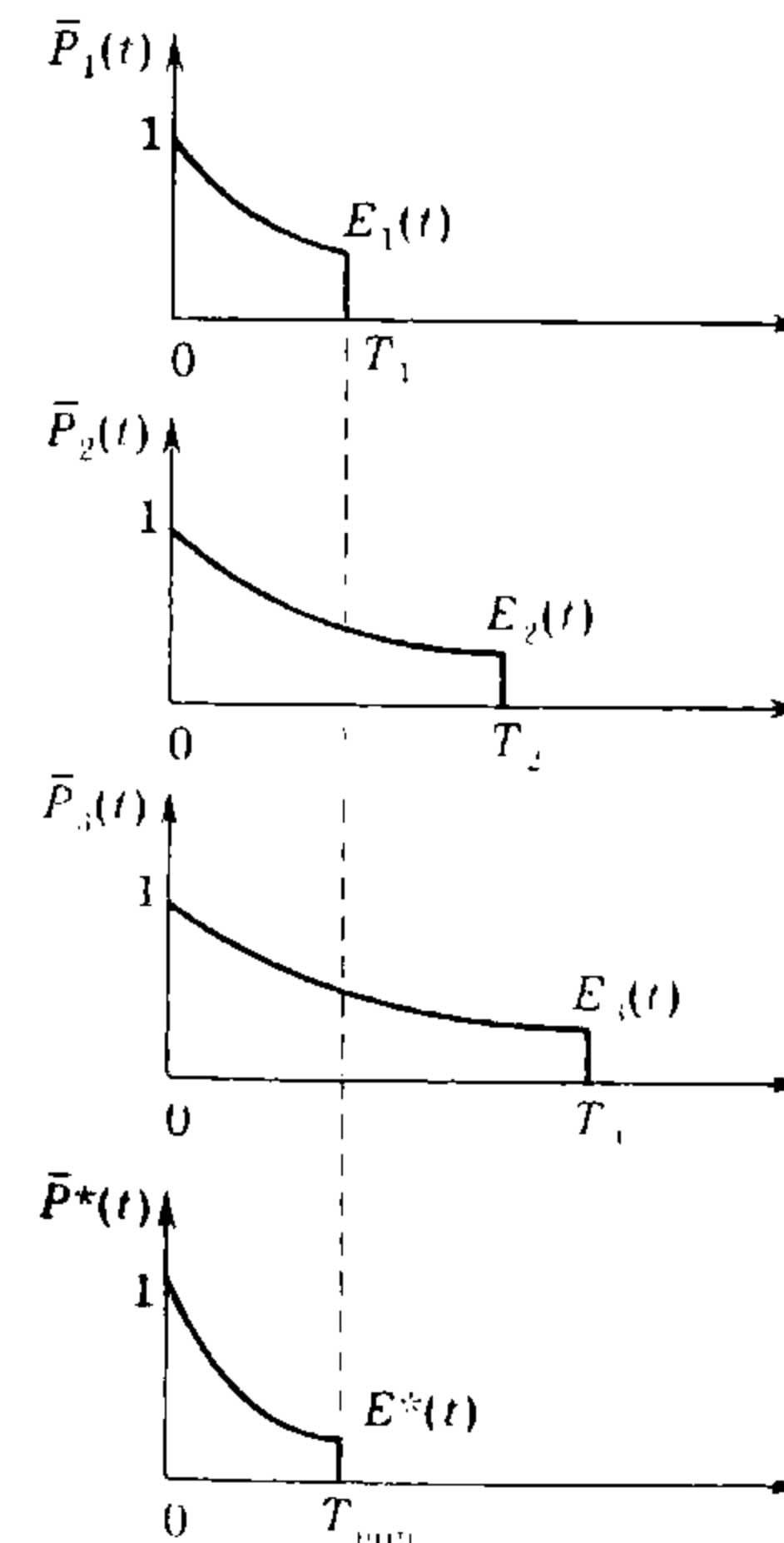


Рис 5.9

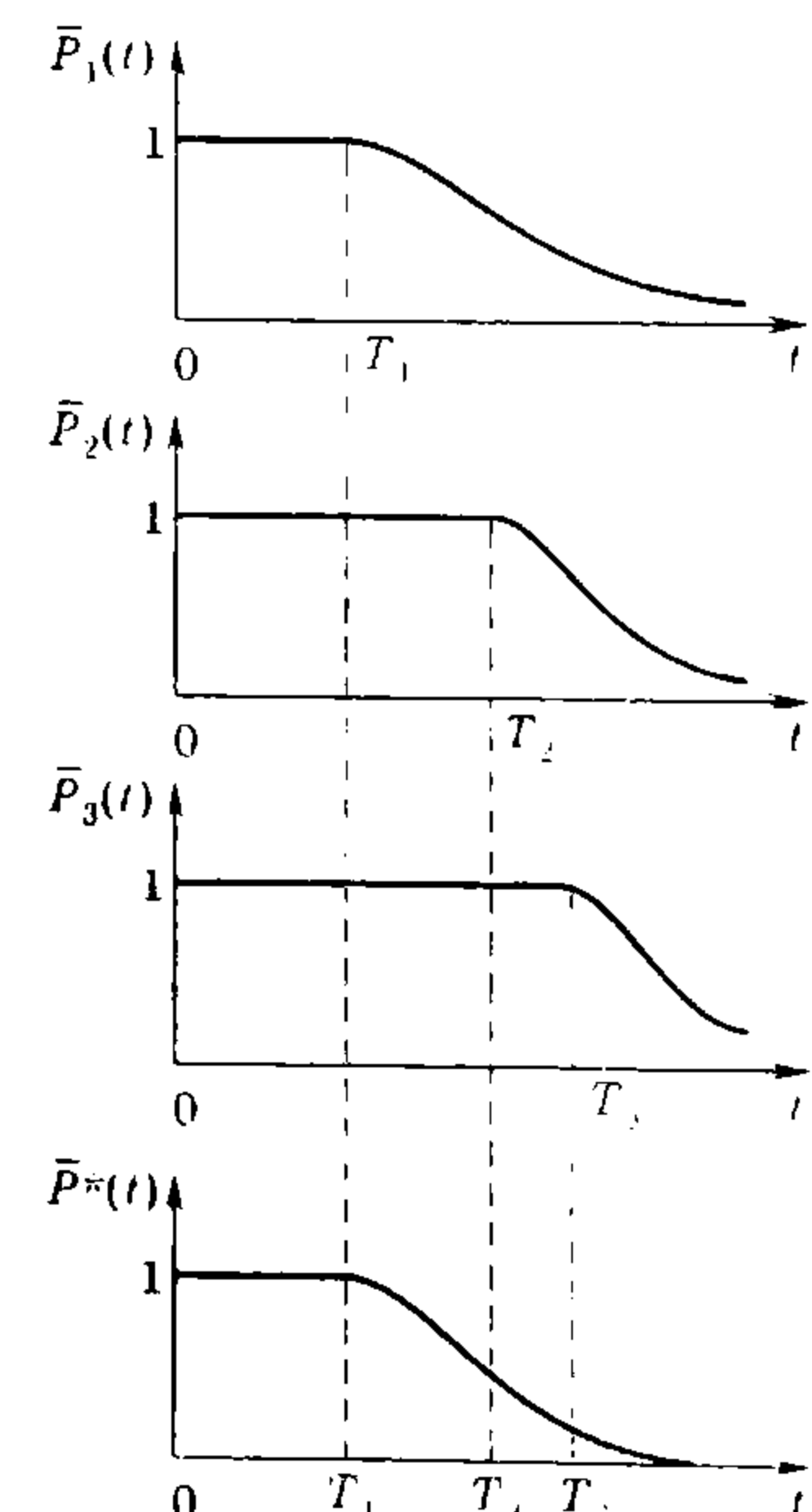


Рис 5.10

Верхняя граница ВБР последовательной системы запишется в виде

$$\bar{P}(t) = \prod_{i=1}^n \bar{P}_i(t). \quad (5.30)$$

После подстановки (5.11) в (5.30) находим, что верхняя граница ВБР последовательной системы из *стареющих* элементов определяется как

$$P(t) \leq \bar{P}(t) = \begin{cases} 1 & \text{для } t \leq T_1, \\ \exp(-\omega_1 t) & \text{для } T_1 < t \leq T_2, \\ \exp[-(\omega_1 t + \omega_2 t)] & \text{для } T_2 < t \leq T_3, \\ \dots & \dots \\ \exp\left(-\sum_{i=1}^n \omega_i t\right) & \text{для } t \geq T_n; \end{cases} \quad (5.31)$$

здесь ω_k зависят от соответствующих моментов усечения экспоненциальных функций и находятся из уравнений типа (5.10).

К сожалению, эта верхняя граница ВБР последовательной системы обладает теми же недостатками, что и (5.11). Еще раз заметим, что на практике представляет больший интерес нижняя граница, дающая гарантированное значение ВБР системы.

Если известны значения $\alpha_i = \lambda_i(0)$ — интенсивности отказов элементов, вычисленные для малых t , — то можно записать выражение верхней границы ВБР системы в виде

$$P_{\text{сист}}(t) \leq \prod_{i=1}^n \exp(-\alpha_i t) = \exp\left(-t \sum_{i=1}^n \alpha_i\right). \quad (5.32)$$

§ 3.3 Среднее время работы до отказа

Верхняя граница. Верхняя граница для среднего времени безотказной работы последовательной системы определяется из свойства последовательной системы: она не может быть надежнее наименее надежного из своих элементов, т. е.

$$P_{\text{сист}}(t) \leq \min_{1 \leq i \leq n} P_i(t). \quad (5.33)$$

Отсюда следует:

$$T_{\text{сист}} \leq \min_{1 \leq i \leq n} \int_0^{\infty} P_i(t) dt = \min_{1 \leq i \leq n} T_i. \quad (5.34)$$

нижняя граница. Для построения нижней границы воспользуемся результатом леммы (5.3.1). Покажем, что замена любого *стареющего* элемента на элемент с экспоненциальным распределением времени до отказа с тем же средним ведет к уменьшению среднего времени безотказной работы последовательной

системы. Допустим, что n -й *стареющий* элемент системы заменен на элемент с экспоненциальным распределением. Вычислим приращение в значении среднего времени:

$$\begin{aligned} \Delta &= \int_0^{\infty} \prod_{i=1}^n P_i(t) dt - \int_0^{\infty} \exp(-t/T_n) \prod_{i=1}^{n-1} P_i(t) dt = \\ &= \int_0^{\infty} [P_n(t) - \exp(-t/T_n)] \prod_{i=1}^{n-1} P_i(t) dt. \end{aligned} \quad (5.35)$$

Поскольку функция $P_n(t)$ пересекает $\exp(-t/T_n)$ один раз, причем сверху, а средние их равны по предположению, функция

$$P_n(t) - \exp(-t/T_n) \quad (5.36)$$

соответствует функции $g(x)$ леммы. В то же самое время функция $\prod_{i=1}^{n-1} P_i(t)$ соответствует убывающей функции $f(x)$ в той же лемме. Таким образом, замена *стареющего* элемента на элемент с экспоненциальным распределением уменьшает значение среднего времени работы до отказа системы в целом, поскольку в (5.35) приращение $\Delta < 0$. Производя подобные подстановки систематически, придем к нижней границе для среднего времени безотказной работы:

$$T_{\text{сист}} \geq \int_0^{\infty} \prod_{i=1}^n \exp\left(-\frac{t}{T_i}\right) dt = \int_0^{\infty} \exp\left(-t \sum_{i=1}^n \frac{1}{T_i}\right) dt = \left(\sum_{i=1}^n \frac{1}{T_i}\right)^{-1}. \quad (5.37)$$

Таким образом, получен окончательный результат: среднее время безотказной работы последовательной системы, состоящей из n независимых *стареющих* элементов с известными средними, заключено в границах:

$$\left(\sum_{i=1}^n \frac{1}{T_i}\right)^{-1} \leq T_{\text{сист}} \leq \min_{1 \leq i \leq n} T_i. \quad (5.38)$$

§ 5.4. ПАРАЛЛЕЛЬНЫЕ СИСТЕМЫ

5.4.1. Вероятность безотказной работы

Верхняя граница. Запишем выражение вероятности отказа параллельной системы:

$$Q_{\text{сист}} = \prod_{i=1}^m Q_i(t). \quad (5.39)$$

Построим нижнюю границу для (5.39), что, естественно, позволит найти верхнюю границу ВБР. Перенумеруем элементы в таком порядке, чтобы выполнялось условие $T_1 < T_2 < \dots < T_m$,

где T_k — среднее время работы до отказа k -го элемента. Заметим, что для элемента с наибольшим средним временем безотказной работы, т. е. m -го элемента, при всех $t < T_m$ выполняется условие $\bar{P}_m(t) = 1$, или, что то же самое, $Q_m(t) = Q_{\text{сист}}(t) = 0$. Для $t \geq T_m$ все $Q_i(t) = 1 - \exp(-\omega_i t)$, где параметры ω_i определены как ω_2 в (5.10). Таким образом, нижняя граница $Q_{\text{сист}}$ вероятности отказа параллельной системы, состоящей из m независимых *стареющих* элементов, имеет вид

$$Q_{\text{сист}}(t) = \begin{cases} 0 & \text{для } t \leq T_m, \\ \prod_{i=1}^m [1 - \exp(-\omega_i t)] & \text{для } t > T_m. \end{cases} \quad (5.40)$$

Здесь по-прежнему ω_i определяются в соответствии с (5.10).

Естественно, что (5.40) соответствует следующей верхней границе ВБР:

$$\bar{P}_{\text{сист}}(t) = \begin{cases} 1 & \text{для } t \leq T_m, \\ 1 - \prod_{i=1}^m [1 - \exp(-\omega_i t)] & \text{для } t > T_m. \end{cases} \quad (5.41)$$

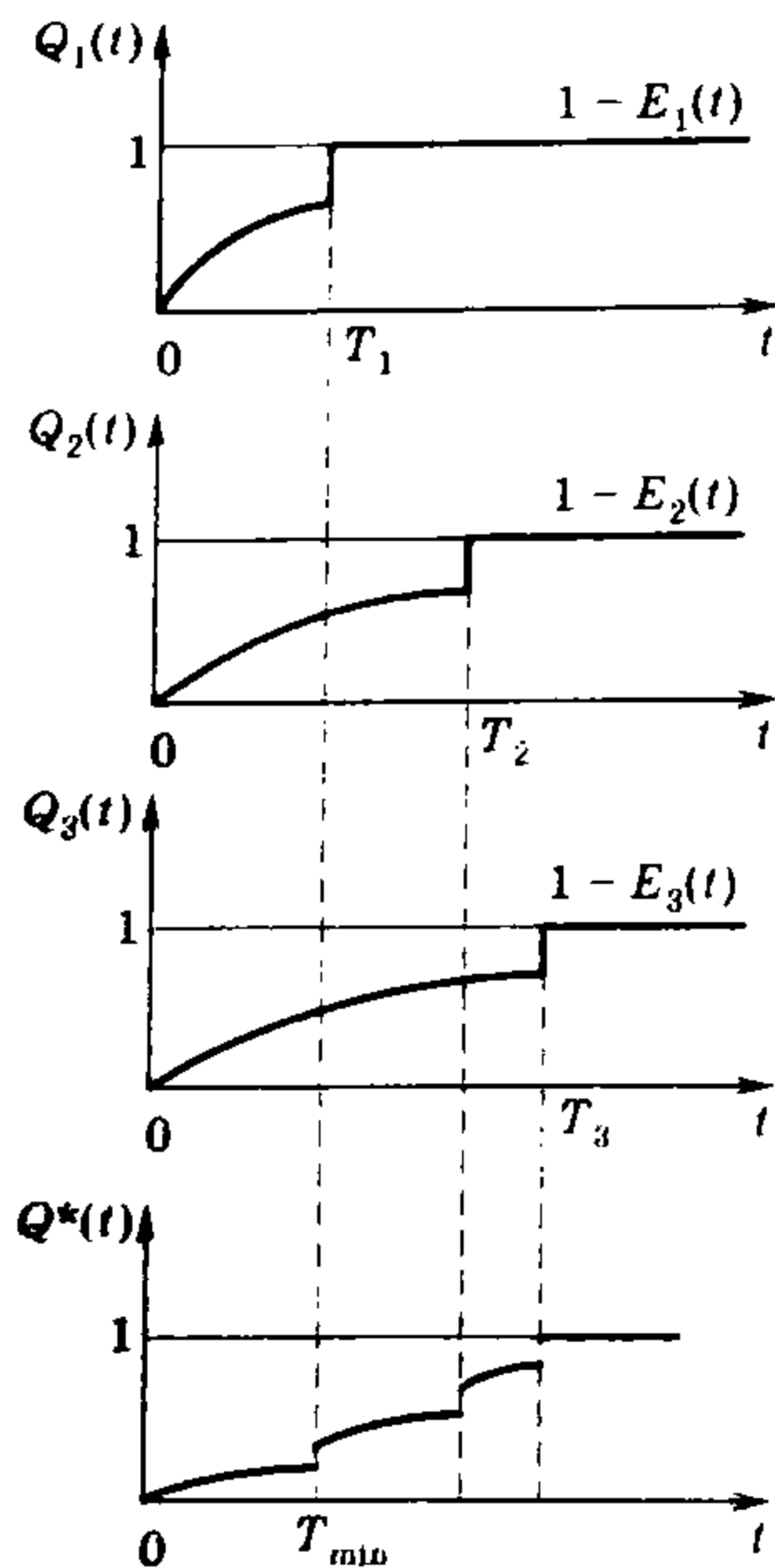


Рис. 5.11

Можно опять отметить, что верхняя оценка для параллельной системы практически не интересна, так как для времен, меньших наибольшего из средних времен безотказной работы элементов, она тривиальна и неинформативна.

нижняя граница. Чтобы найти нижнюю границу ВБР параллельной системы, найдем сначала верхнюю границу для вероятности отказа. Обратимся опять к формуле (5.39) и воспользуемся рис. 5.11, который является, в некотором смысле, зеркальным к рис. 5.9.

На этом рисунке дан пример системы, состоящей из трех элементов. Первые три графика соответствуют экспоненциальным функциям, усеченным в точках T_1 , T_2 и T_3 . Результирующая функция, равная произведению трех верхних, представлена

на нижнем графике. Она и представляет собой нижнюю границу для вероятности отказа параллельной системы. Из рисунка видно, что при $t > T_3 = T_{\text{max}}$ для вероятности отказа системы имеем $Q_{\text{сист}}(t) = 1$, поэтому нижняя граница ВБР системы на том же интервале времени $\bar{P}_{\text{сист}}(t) = 0$. Из того же рисунка следует, что на интервале времени $t \leq T_1$ верхняя граница вероятности отказа равна $\bar{Q}_{\text{сист}} = [1 - \exp(-t/T_1)] \cdot [1 - \exp(-t/T_2)] \cdot [1 - \exp(-t/T_3)]$, на интервале $T_1 < t \leq T_2$ эта граница равна $\bar{Q}_{\text{сист}} = [1 - \exp(-t/T_2)] \times [1 - \exp(-t/T_3)]$ и, наконец, на интервале $T_2 < t \leq T_3$ она равна $\bar{Q}_{\text{сист}} = 1 - \exp(-t/T_3)$.

Проведя аналогичные рассуждения для общего случая, получаем

$$\bar{Q}_{\text{сист}}(t) = \begin{cases} \prod_{i=1}^m [1 - \exp(-t/T_i)] & \text{для } t \leq T_1, \\ \prod_{i=2}^m [1 - \exp(-t/T_i)] & \text{для } T_1 < t \leq T_2, \\ \dots & \dots \\ 1 & \text{для } t > T_m. \end{cases} \quad (5.42)$$

Из этого следует выражение нижней границы ВБР параллельной системы:

$$P_{\text{сист}}(t) = \begin{cases} 1 - \prod_{i=1}^m [1 - \exp(-t/T_i)] & \text{для } t \leq T_1, \\ 1 - \prod_{i=2}^m [1 - \exp(-t/T_i)] & \text{для } T_1 < t \leq T_2, \\ \dots & \dots \\ 0 & \text{для } t > T_m. \end{cases} \quad (5.43)$$

Заметим, что наиболее существенные в инженерной практике нижние (т. е. гарантированные) границы ВБР как последовательной, так и параллельной систем достаточно информативны.

5.4.2. Среднее время работы до отказа

нижняя граница. Эта граница для среднего времени безотказной работы определяется тем, что ВБР параллельной системы не ниже ВБР любого, даже самого надежного, ее элемента, т. е.

$$T_{\text{сист}} = \int_0^{\infty} \left[1 - \prod_{i=1}^m Q_i(t) \right] dt \geq \max_{1 \leq i \leq m} \int_0^{\infty} [1 - Q_i(t)] dt = \\ = \max_{1 \leq i \leq m} \int_0^{\infty} P_i(t) dt = T_{\text{max}}. \quad (5.44)$$

■ ВЕРХНЯЯ ГРАНИЦА. При построении верхней границы среднего времени безотказной работы можно воспользоваться леммой, сформулированной в 5.3.1. Опять перенумеруем элементы так, чтобы выполнялось условие $T_1 < T_2 < \dots < T_m$. Запишем выражение для среднего времени безотказной работы в виде

$$T_{\text{сист}} = \int_0^{\infty} \left[1 - \prod_{i=1}^n Q_i(t) \right] dt. \quad (5.45)$$

Заменим n -й *стареющий* элемент на экспоненциальный с тем же средним и проверим знак разности

$$\begin{aligned} \Delta &= \int_0^{\infty} \left[1 - \prod_{i=1}^n Q_i(t) \right] dt - \int_0^{\infty} \left[1 - [1 - \exp(-t/T_n)] \cdot \prod_{i=1}^{n-1} Q_i(t) \right] dt = \\ &= \int_0^{\infty} \left\{ \left[1 - \prod_{i=1}^n Q_i(t) \right] - \left[1 - [1 - \exp(-t/T_n)] \cdot \prod_{i=1}^{n-1} Q_i(t) \right] \right\} dt = \\ &= \int_0^{\infty} \left\{ [1 - \exp(-t/T_n)] \cdot \prod_{i=1}^{n-1} Q_i(t) - \prod_{i=1}^n Q_i(t) \right\} dt = \\ &= \int_0^{\infty} \prod_{i=1}^{n-1} Q_i(t) \cdot \left\{ [1 - \exp(-t/T_n)] - Q_n(t) \right\} dt = \\ &= \int_0^{\infty} \prod_{i=1}^{n-1} Q_i(t) \cdot \left\{ [P_n(t) - \exp(-t/T_n)] \right\} dt. \end{aligned} \quad (5.46)$$

Подынтегральная функция представляет собой произведение двух сомножителей: первый есть произведение вероятностей отказа, т. е. возрастающая функция времени, а второй — разность, удовлетворяющая условиям леммы на функцию $g(t)$. В соответствии с леммой, величина Δ отрицательна, это значит, что замена *стареющего* элемента на элемент с соответствующим экспоненциальным распределением увеличивает среднее время безотказной работы системы.

Используя подобную подстановку систематически, придем к окончательному результату:

$$\begin{aligned} T_{\text{сист}} &\leq \int_0^{\infty} \left[1 - \prod_{i=1}^m (1 - \exp(-t/T_i)) \right] dt = \\ &= \int_0^{\infty} \sum_{i=1}^m \exp(-t/T_i) dt - \int_0^{\infty} \sum_{1 \leq i < j \leq m} \exp(-t/T_i) \exp(-t/T_j) dt + \dots + \\ &\quad + (-1)^{m+1} \int_0^{\infty} \sum_{i=1}^m \exp(-t/T_i) dt, \end{aligned} \quad (5.47)$$

что в результате дает желаемую оценку после взятия интегралов.

Итак, среднее время безотказной работы параллельной системы, состоящей из m независимых *стареющих* элементов, имеет следующие границы:

$$\begin{aligned} \max_{1 \leq i \leq m} T_i &\leq T_{\text{сист}} \leq \sum_{i=1}^m T_i - \sum_{1 \leq i < j \leq m} (T_i^{-1} + T_j^{-1})^{-1} + \dots + \\ &\quad + (-1)^{m+1} \left(\sum_{i=1}^m T_i^{-1} \right)^{-1}. \end{aligned} \quad (5.48)$$

§ 5.5. ГРАНИЦЫ ДЛЯ КОЭФФИЦИЕНТА ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ

Как уже было определено, коэффициент интервальной готовности есть произведение двух вероятностей: вероятности заставить систему в состоянии работоспособности в произвольный момент времени (коэффициент готовности) и вероятности безотказно проработать с этого произвольного момента в течение требуемого времени t_0 . Стационарный коэффициент готовности не зависит от распределения времени работы между отказами, но такая зависимость имеет место для вероятности безотказно проработать с произвольного момента. Поэтому прежде всего необходимо найти ее границы.

Из теории восстановления известно, что стационарное распределение остаточного времени $R(t)$ можно характеризовать выражением

$$R(t_0) = \frac{1}{T} \int_{t_0}^{\infty} P(t) dt, \quad (5.49)$$

в котором $P(t)$ есть ВБР, т. е. функция дополнительная к функции распределения, T есть среднее, а t_0 — заданное время выполнения системой своих операций.

Снова воспользуемся леммой из 5.3.1. Ее легко модифицировать, введя в качестве $f(x)$ пороговую функцию, имеющую скачок в точке t_0 . Пусть функция $g(x)$ абсолютно интегрируема, $g(x) \geq 0$

при $x < a$ и $g(x) \leq 0$ при $x \geq a$, а также $\int_0^{\infty} g(x) dx = 0$. Тогда имеем

$$\int_{t_0}^{\infty} g(x) dx < 0. \quad (5.50)$$

Рассмотрим опять функции: $P(t)$ для *стареющего*, $G(t)$ — вырожденного и $E(t)$ — экспоненциального распределений. Возь-

мем сначала функцию $g_1(x) = P(t) - G(t)$ и применим к ней модифицированную лемму:

$$\int_{t_0}^{\infty} [P(t) - G(t)] dx < 0 \Rightarrow \int_{t_0}^{\infty} P(t) dx < \int_{t_0}^{\infty} G(t) dx. \quad (5.51)$$

Затем возьмем функцию $g_2(x) = P(t) - G(t)$, для которой

$$\int_{t_0}^{\infty} [E(t) - P(t)] dx < 0 \Rightarrow \int_{t_0}^{\infty} E(t) dx < \int_{t_0}^{\infty} P(t) dx. \quad (5.52)$$

Из (5.41) и (5.42) следует:

$$\int_{t_0}^{\infty} G(t) dx < \int_{t_0}^{\infty} P(t) dx < \int_{t_0}^{\infty} E(t) dx. \quad (5.53)$$

Поскольку $\int_{t_0}^{\infty} G(t) dt = T - t_0$ и $\int_{t_0}^{\infty} E(t) dt = \int_{t_0}^{\infty} \exp(-t/T) dt$, после подстановки этих значений в (5.53) получаем

$$T - t_0 < \int_{t_0}^{\infty} P(t) dt < T \exp(t_0/T). \quad (5.54)$$

В результате можно записать окончательно неравенства для (5.49) в виде

$$1 - t_0/T < R(t_0) < \exp(-t_0/T). \quad (5.55)$$

При $t_0 > T$ левая часть неравенства обращается в нуль.

Отметим, что полученные граничные оценки стационарного распределения остаточного времени работы являются очень хорошими особенно для высоконадежных систем; при этом справедливо приближенное выражение

$$R(t_0) \approx (1 - t_0/T). \quad (5.56)$$

Стационарный коэффициент интервальной готовности составляет $R^*(t) = KR(t)$; здесь K — стационарный коэффициент готовности. Это позволяет для высоконадежных систем написать приближенное выражение в виде

$$R^*(t_0) \approx \left(1 - \frac{t_0}{T}\right) \left(1 - \frac{\tau}{T}\right) \approx 1 - \frac{t_0 + \tau}{T}; \quad (5.57)$$

в (5.57) через τ обозначено среднее время простоя.

По своему качеству данные верхняя и нижняя границы коэффициента интервальной готовности, равно как и приближение, полученное с их помощью, являются наилучшими и широко используются на практике.

Глава 6

Эвристические методы расчета надежности

Разработаны строгие асимптотические методы анализа высоконадежных систем. Однако эти методы зачастую слишком сложны, а потому не часто используются в повседневной инженерной практике. Предлагаются эвристические методы расчета надежности, которые дают приемлемые результаты.

§ 6.1. ОДИНОЧНЫЙ ЭЛЕМЕНТ

Процесс функционирования одиночного восстанавливаемого элемента может быть представлен в виде альтернирующего процесса: чередующихся интервалов работы ξ и простоя η , имеющих произвольные распределения $F(t)$ и $G(t)$, соответственно. На практике обычно рассматриваются высоконадежные восстанавливаемые элементы, для которых выполняется условие

$$\frac{E(\eta)}{E(\xi)} \ll 1. \quad (6.1)$$

В этом случае альтернирующий процесс, описывающий функционирование восстанавливаемого элемента, может быть представлен как поток очень узких «импульсов» интервалов простоя η_k , разделенных большими интервалами наработки между отказами, ξ_k (рис. 6.1).

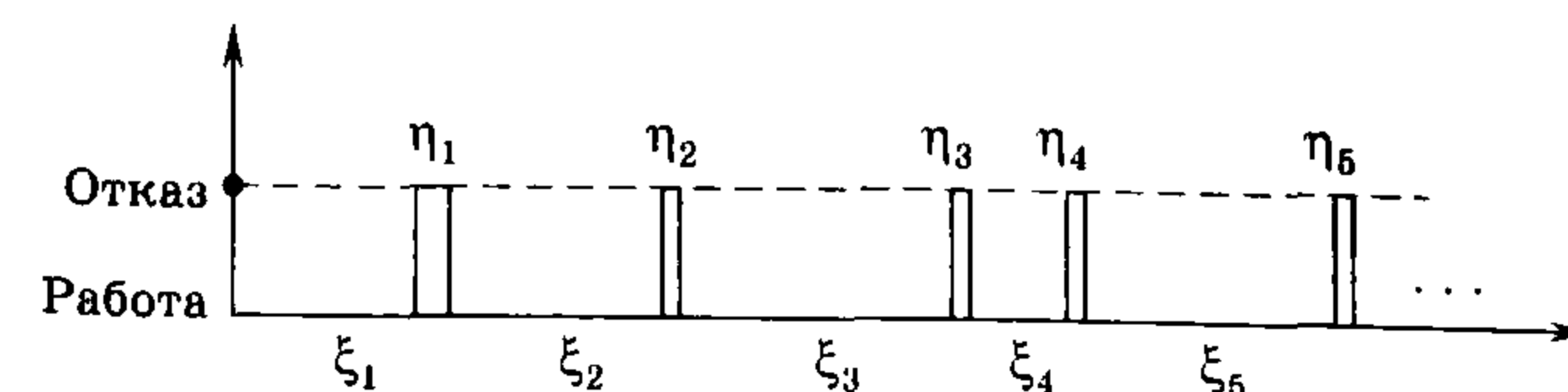


Рис. 6.1

Обозначим средние рассматриваемых случайных величин через $T = E\{\xi\}$ и $\tau = E\{\eta\}$ и запишем основные показатели надежности восстанавливаемого элемента.

Вероятность безотказной работы элемента на интервале времени $[0, t_0]$ равна

$$P(t_0) = 1 - F(t_0). \quad (6.2)$$

Коэффициент готовности K для высоконадежного восстанавливаемого элемента можно записать приближенно в виде

$$K = \frac{T}{T + \tau} \approx 1 - \frac{\tau}{T}. \quad (6.3)$$

Коэффициент интервальной готовности, воспользовавшись результатом (5.57), можно представить следующим образом:

$$R(t_0) \approx K \exp\left(-\frac{t_0}{T}\right) \approx \left(1 - \frac{\tau}{T}\right) \cdot \left(1 - \frac{t_0}{T}\right), \quad (6.4)$$

или окончательно

$$R(t_0) \approx 1 - \frac{\tau + t_0}{T}. \quad (6.5)$$

Итак, для одиночного элемента никаких новых результатов не получено, однако дано описание функционирования, удобное для дальнейших выводов с использованием эвристического метода анализа систем с восстановлением.

§ 6.2. ПРЕДЕЛЬНАЯ ТЕОРЕМА О СУПЕРПОЗИЦИИ ПОТОКОВ

6.2.1. Общее описание регенерирующих точечных случайных процессов

В практических задачах интервал времени между отказами обычно оказывается во много раз больше интервала восстановления работоспособности, поэтому для многих задач надежности, связанных с анализом восстанавливаемых систем, можно условно рассматривать отказы как точки на временной оси. В этом случае удастся эффективно применять математический аппарат теории случайных точечных процессов, в частности предельную теорему о суперпозиции потоков*.

* Основная предельная теорема была доказана русскими математиками А. Я. Хинчиным и Г. А. Ососковым. Обобщение их теоремы было позднее получено литовским математиком Б. И. Григелионисом.



Рис. 6.2

Случайным точечным процессом называется поток точечных событий, которые происходят через случайные интервалы времени.

В общем случае эти интервалы могут быть зависимыми, однако далее мы будем рассматривать только регенерирующие точечные процессы*, в которых случайные интервалы между событиями взаимно независимы и одинаково распределены. Такой тип случайных точечных процессов наиболее интенсивно используется в теории надежности.

Ради удобства последующего изложения будем использовать термин *отказ*, а не *событие*.

Пусть одновременно наблюдается несколько случайных точечных потоков. Суперпозицией точечных потоков называется новый случайный точечный поток, который включает все события исходных потоков. Процедура суперпозиции может быть пояснена графически с помощью рис. 6.2.

6.2.2. Суперпозиция пуассоновских потоков

Рассмотрим сначала суперпозицию конечного числа n пуассоновских потоков с параметрами, соответственно, $\lambda_1, \lambda_2, \dots, \lambda_n$ в произвольный момент времени t . Для k -го пуассоновского потока распределение времени от произвольного момента до очередного отказа имеет экспоненциальное распределение с параметром λ_k , поскольку случайное остаточное время работы, ζ_k , имеет то же

* В отечественной математической литературе регенерирующие процессы чаще называются процессами восстановления. Мы старались в книге по надежности избежать терминологического конфликта между случайным точечным процессом восстановления и чисто физическим процессом восстановления, состоящим в проведении восстановительных работ с отказавшим объектом. В связи с этим в дальнейшем используется термин *регенерирующий процесс*.

экспоненциальное распределение, что и время работы до отказа. Появление первого отказа в результирующем потоке соответствует появлению самого раннего отказа из всех n образующих его потоков. Как известно, распределение минимума экспоненциальных случайных величин является экспоненциальным:

$$P\left\{\min_{1 \leq k \leq n} \zeta_k > x\right\} = P\{\zeta_1 > x, \zeta_2 > x, \dots, \zeta_n > x\} = \prod_{k=1}^n P\{\zeta_k > x\} = \prod_{k=1}^n e^{-\lambda_k x}. \quad (6.6)$$

В результате можно записать

$$P\left\{\min_{1 \leq k \leq n} \zeta_k > x\right\} = \exp\left(-x \prod_{k=1}^n \lambda_k\right). \quad (6.7)$$

Таким образом, суперпозиция произвольного числа пуассоновских потоков порождает пуассоновский же поток.

6.2.3. Суперпозиция произвольных регенерирующих потоков

ТЕОРЕМА. Суперпозиция произвольных регенерирующих точечных процессов при достаточно слабых условиях сходится к пуассоновскому потоку.

Строгое доказательство этой теоремы является достаточно сложным и опускается.

На интуитивном уровне этот факт может быть объяснен следующим образом. Если потоки восстановления имеют непрерывные распределения, то их суперпозиция не может приобрести свойство неординарности, т. е. совпадение событий или образование *точек конденсации* невозможно. Суперпозиция стационарных исходных процессов не может привести к образованию нестационарного результирующего процесса.

Наконец, если участвующие в суперпозиции потоки независимы и вносят в него *равномерно малый вклад*, то при большом числе суммируемых потоков результирующий состоит из практически независимых событий.

Более строгое рассуждение для суперпозиции произвольных потоков, образованных *стареющими* распределениями, может быть проведено следующим образом. Покажем, что время от произвольного момента времени в стационарном режиме до первого отказа суперпозиции потоков имеет экспоненциальное распределение. Тем самым будет доказано, что соответствующий результи-

рующий поток является пуассоновским. В главе 5 было показано, что остаточное время работы ζ для восстанавливаемого элемента имеет при малых t распределение, близкое к экспоненциальному:

$$1 - \frac{t}{T} \leq P\{\zeta \geq t\} \leq \exp\left(-\frac{t}{T}\right). \quad (6.8)$$

Это приводит нас к следующим границам для суперпозиции потоков отказов:

$$1 - \frac{t}{T^*} \leq P\left\{\min_{1 \leq k \leq n} \zeta_k > x\right\} = \prod_{k=1}^n P\{\zeta_k > x\} \leq \prod_{k=1}^n \exp\left(-\frac{t}{T_k}\right) = \exp\left(-t \sum_{k=1}^n \frac{1}{T_k}\right) = \exp\left(-\frac{t}{T^*}\right); \quad (6.9)$$

здесь через T^* обозначено $T^* = \left[\sum_{i=1}^n (1/T_i)\right]^{-1}$.

§ 6.3. ВЫСОКОНАДЕЖНАЯ ПОСЛЕДОВАТЕЛЬНАЯ СИСТЕМА

Рассмотрим высоконадежную восстанавливаемую последовательную систему из n независимых элементов, для которой выполняется условие

$$\max_{1 \leq i \leq n} \frac{\tau_i}{T_i} \ll \frac{1}{n}, \quad (6.10)$$

в котором T_i — среднее время безотказной работы i -го элемента, τ_i — среднее время восстановления системы после отказа i -го элемента.

Процесс, описывающий функционирование последовательной системы, представляет собой суперпозицию потоков *импульсов простоя* ее элементов. Пример для $n = 3$ приведен на рис. 6.3.

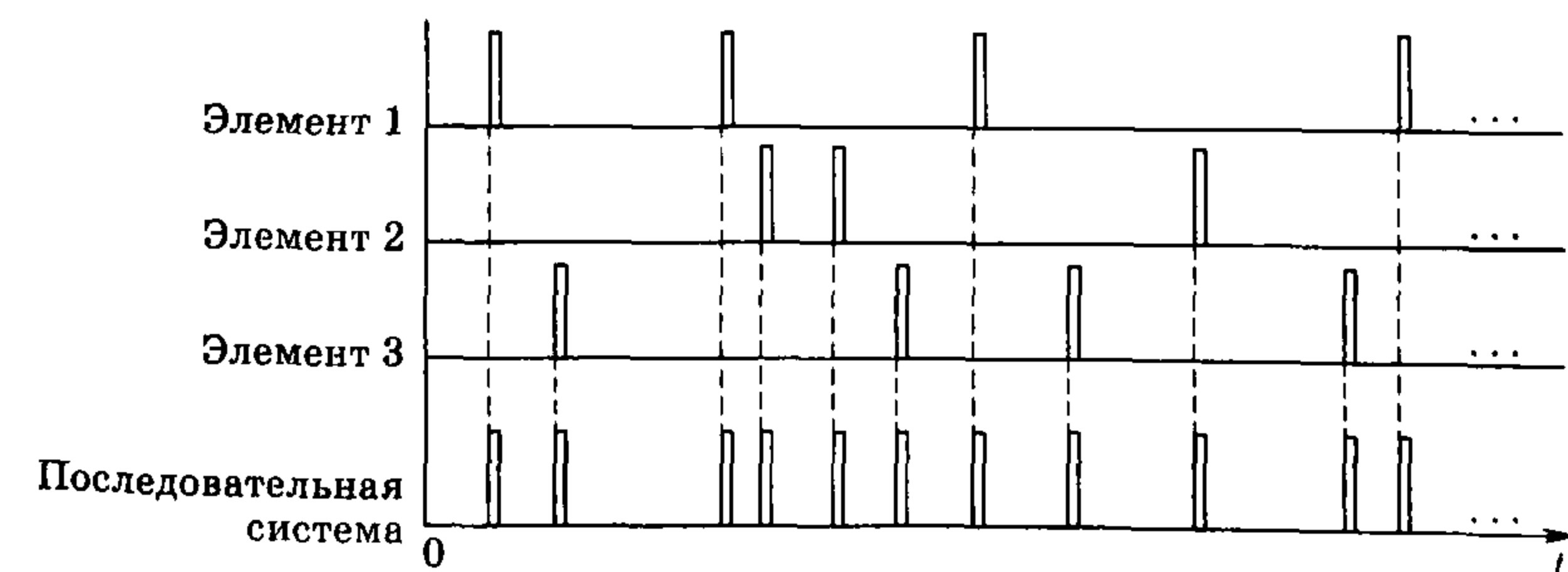


Рис. 6.3

Поскольку потоки отказов элементов предполагаются независимыми, а «импульсы простоя» предполагаются очень короткими, это позволяет пренебречь возможностью их перекрытия. Согласно предельной теореме суперпозиции случайных точечных потоков при больших n поток отказов последовательной системы будет близким к пуассоновскому с параметром

$$\lambda_{\text{сист}} = \sum_{i=1}^n \frac{1}{T_i}. \quad (6.11)$$

Время простоя системы будет иметь гиперэкспоненциальное распределение (см. Приложение 2). Его среднее значение может быть вычислено как средневзвешенное:

$$\tau_{\text{сист}} = \frac{1}{\lambda_{\text{сист}}} \sum_{i=1}^n \lambda_i \tau_i. \quad (6.12)$$

ВЕРЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ для последовательной системы может быть приближенно представлена в виде

$$P(t_0) \approx \exp\left(-t_0 \sum_{i=1}^n \frac{1}{T_i}\right). \quad (6.13)$$

КОЭФФИЦИЕНТ ГОТОВНОСТИ K для последовательной системы с восстановлением равен

$$K = \frac{1}{1 + \lambda_{\text{сист}} \tau_{\text{сист}}} \approx 1 - \sum_{i=1}^n \lambda_i \tau_i. \quad (6.14)$$

КОЭФФИЦИЕНТ ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ опять, воспользовавшись (5.7), можно записать следующим образом:

$$R(t_0) \approx \left(1 - \sum_{i=1}^n \lambda_i \tau_i\right) \cdot \left(1 - t_0 \sum_{i=1}^n \lambda_i\right) \approx \left(1 - \sum_{i=1}^n \lambda_i (\tau_i + t_0)\right). \quad (6.15)$$

ПРИМЕР 6.1. Последовательная система состоит из элементов трех типов: a , b и c . Исходные данные сведены в табл. 6.1. Длительность заданной операции системы равна $t_0 = 2$ ч. Требуется найти основные показатели надежности рассматриваемой последовательной системы.

Решение. Интенсивность отказов системы:

$$\lambda_{\text{сист}} = 50 \cdot 5 \cdot 10^{-5} + 50 \cdot 2.5 \cdot 10^{-5} + 100 \cdot 2 \cdot 10^{-5} = 4.75 \cdot 10^{-3} [1/\text{ч}].$$

Среднее время простоя системы:

$$\tau_{\text{сист}} = (4.75 \cdot 10^{-3})^{-1} \cdot (50 \cdot 5 \cdot 10^{-5} \cdot 0.5 + 50 \cdot 2.5 \cdot 10^{-5} \cdot 1 + 100 \cdot 2 \cdot 10^{-5} \cdot 0.5) = 0.737 [\text{ч}].$$

Среднее время безотказной работы системы:

$$T_{\text{сист}} = (4.75 \cdot 10^{-3})^{-1} \approx 210.5 [\text{ч}].$$

Таблица 6.1

Типы элементов	a	b	c
Число элементов n_i	50	50	100
Среднее время безотказной работы T_i , ч	20 000	40 000	50 000
Среднее время восстановления τ_i , ч	0.5	1	0.5

Вероятность безотказной работы для операции длительностью 2 ч:

$$P(2) \approx 1 - 2 \cdot 4.75 \cdot 10^{-3} = 0.9905.$$

Коэффициент готовности:

$$K \approx 1 - (50 \cdot 5 \cdot 10^{-5} \cdot 0.5 + 50 \cdot 2.5 \cdot 10^{-5} \cdot 1 + 100 \cdot 2 \cdot 10^{-5} \cdot 0.5) = 0.9965.$$

Коэффициент интервальной готовности для операции длительностью 2 ч:

$$R(2) \approx 0.9905 \cdot 0.9965 \approx 0.996.$$

§ 6.4. ПРЕДЕЛЬНАЯ ТЕОРЕМА О РАЗРЕЖЕНИИ ПОТОКОВ

При анализе надежности резервированных систем весьма эффективно используется теорема о разрежении потоков*.

Рассмотрим реализацию произвольного случайного точечного потока. Процедура разрежения случайного потока событий производится по схеме испытаний Бернулли: с вероятностью p событие исключается из потока, а с вероятностью q сохраняется (рис. 6.4). Подобная процедура может быть затем еще неоднократно применена к оставшемуся потоку событий, причем вероятности p и q могут отличаться от предыдущих. Главное требование, чтобы $\prod_{k=1}^N q_k \rightarrow 0$ при $N \rightarrow \infty$, где N — число операций разрежения.

Если применить процедуру разрежения с произвольным, но повсюду постоянным p к пуассоновскому потоку с параметром λ , то можно заметить, что результирующий поток не меняет характеристических свойств исходного потока: поток остается стационар-

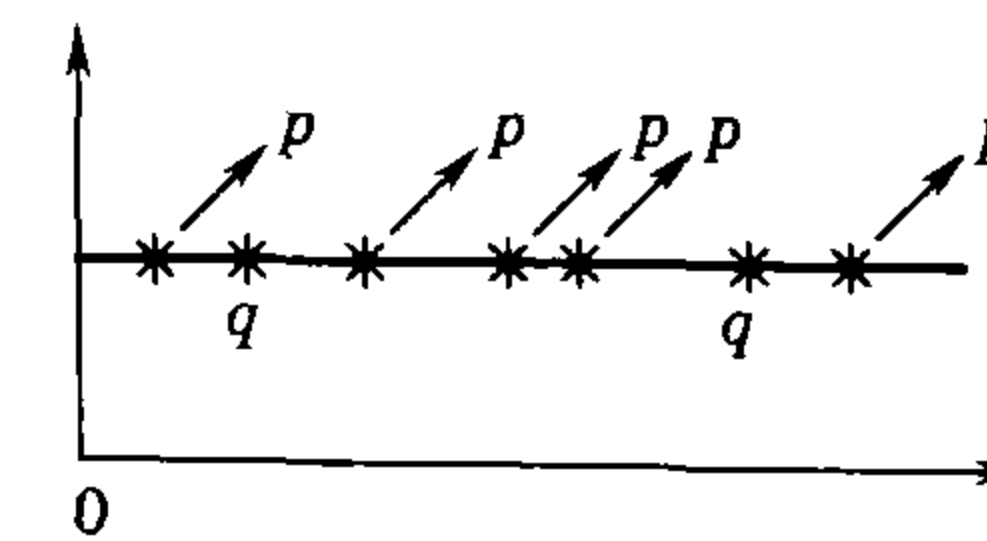


Рис. 6.4

* Основная предельная теорема была доказана венгерским математиком А. Реньи в 40-х годах прошлого столетия. Затем результат был обобщен современным отечественным математиком Ю. К. Беляевым.

ным (вероятности p не меняются во времени), без последствия (исключение какого-либо события не влияет на вероятность исключения других событий) и ординарным (события из потока только лишь исключаются). Это в свою очередь означает, что вновь образованный поток остается пуассоновским, а его интенсивность становится равной λq .

Применим операцию разрежения к произвольному случайному процессу восстановления последовательно n раз. Результирующая вероятность того, что событие сохранится в потоке после n разрежений, есть

$$Q_n = q_1 \cdot q_2 \cdot \dots \cdot q_n.$$

ТЕОРЕМА РАЗРЕЖЕНИЯ. Если $Q_n = q_1 \cdot q_2 \cdot \dots \cdot q_n \rightarrow 0$ при $n \rightarrow \infty$, то результирующий поток после соответствующей нормировки будет сходиться к пуассоновскому.

В теории надежности эта теорема имеет большое значение. По сути дела она постулирует, что распределение суммы случайного числа случайных величин сходится в пределе к экспоненциальной случайной величине, если случайное число слагаемых в сумме имеет геометрическое распределение.

Нормирование вновь получаемой случайной величины совершенно необходимо вследствие того, что при $Q_n \rightarrow 0$ математическое ожидание интервала между отказами стремится к бесконечности, т. е. распределение становится несобственным. Умножая случайный интервал между событиями на $(Q_n)^{-1}$, мы сохраняем математическое ожидание постоянным, равным начальному значению.

Доказательство. Рассмотрим произвольный процесс восстановления с ведущей функцией $F(t)$. Обозначим соответствующую этому распределению случайную величину через X . Введем нормированную случайную величину X_v , равную сумме случайного числа v слагаемых:

$$X_v = \frac{\sum_{i=1}^v X_i}{E\left\{\sum_{i=1}^v X_i\right\}}. \quad (6.16)$$

Математическое ожидание этой новой случайной величины равно единице: $E\{X_v\} = 1$. По тождеству Вальда для математического ожидания суммы случайного числа случайных переменных можно записать:

$$E\left\{\sum_{i=1}^v X_i\right\} = E\{v\} E\{X\}. \quad (6.17)$$

Величину $E\{X\}$ также можно без потери общности положить равной единице. Поскольку v имеет геометрическое распределение с параметром q , значение $E\{v\} = 1/q$. Теперь (6.16) может быть записано в виде

$$X_v = q \sum_{i=1}^v X_i. \quad (6.18)$$

Запишем преобразование Лапласа для плотности распределения $f_v(t)$ случайной величины X_v :

$$\begin{aligned} \varphi_v(s) &= E\left\{e^{-sX_v}\right\} = E\left\{\exp\left(-qs \sum_{i=1}^v X_i\right)\right\} = \\ &= \sum_{k=1}^{\infty} p^{k-1} q \cdot \exp\left(-qs \sum_{i=1}^v X_i\right). \end{aligned} \quad (6.19)$$

Заметим, что

$$\Pr\{v=k\} = qp^{k-1}. \quad (6.20)$$

Тогда можно переписать:

$$\varphi_v(s) = \sum_{k=1}^{\infty} p^{k-1} q [\varphi(sq)]^k = q\varphi(sq) \sum_{k=0}^{\infty} [p\varphi(sq)]^k = \frac{q\varphi(sq)}{1 - p\varphi(sq)}, \quad (6.21)$$

где $\varphi(s) \Leftrightarrow F(t)$. Проведем теперь ряд формальных преобразований:

$$\begin{aligned} \varphi_v(s) &= \frac{q\varphi(sq)}{1 - (1-q)\varphi(sq)} = \\ &= \frac{q\varphi(sq)}{1 - \varphi(sq) + q\varphi(sq)} = \frac{\varphi(sq)}{s \frac{1 - \varphi(sq)}{sq} + \varphi(sq)}. \end{aligned} \quad (6.22)$$

Заметим, что $\varphi(0) = 1$, поэтому

$$\lim_{q \rightarrow 0} \varphi(sq) = 1. \quad (6.23)$$

В результате можно записать

$$\lim_{q \rightarrow 0} \frac{1 - \varphi(sq)}{sq} = \lim_{q \rightarrow 0} \frac{\varphi(0) - \varphi(sq)}{sq} = -\frac{d}{ds} \varphi(s) \Big|_{s=0}. \quad (6.24)$$

Поскольку для преобразования Лапласа от функции распределения справедливо равенство $\frac{d}{ds} \varphi(s) \Big|_{s=0} = -E\{X\}$, учитывая, что

$E\{X\} = 1$, из (6.22) получаем окончательно преобразование Лапласа в виде

$$\varphi_v(s) = \frac{1}{1+s}, \quad (6.25)$$

соответствующее экспоненциальной функции с параметром $\lambda = 1$. Напомним, что без потери общности ранее мы полагали $E\{X\} = 1$.

Из экспоненциальности распределения между отказами в разреженном потоке следует, что предельный поток является пуассоновским, что и требовалось доказать.

§ 6.5. ВЫСОКОНАДЕЖНОЕ ХОЛОДНОЕ ДУБЛИРОВАНИЕ

6.5.1. Вероятность разрежения для потока отказов

Рассмотрим систему, состоящую из одного основного (рабочего) элемента и одного резервного, находящегося в ненагруженном режиме. Элементы предполагаются идентичными. При *холодном* резерве после отказа основного элемента включается запасной. Если и этот элемент отказывает в течение времени восстановления ранее отказавшего, то это приводит к отказу системы. Если же восстановление отказавшего элемента происходит безболезненно, то система «не замечает» происшедшего отказа элемента, т. е. реализуется процедура разрежения потока отказов.

Обозначим случайное время работы элемента до отказа через ξ , а случайное время восстановления элемента через η . Пусть $F(t)$ есть распределение случайной величины ξ , $g(t)$ — плотность распределения случайной величины η , а их средние значения равны T и τ соответственно. Тогда полная вероятность того, что отказ элемента перерастет в отказ системы, может быть записана как

$$q = P\{\xi < \eta\} = \int_0^{\infty} F(t)g(t)dt. \quad (6.26)$$

Если $F(t)$ есть экспоненциальное распределение с параметром $\lambda = 1/T$, и при этом система высоконадежна, т. е. $\lambda\tau \ll 1$, то можно записать приближенное выражение, используя разложение экспоненты в ряд:

$$\begin{aligned} q &= \int_0^{\infty} [1 - \exp(-\lambda t)] g(t)dt = \\ &= \int_0^{\infty} \left(\lambda t - \frac{(\lambda t)^2}{2} + \dots \right) g(t)dt \approx \lambda \int_0^{\infty} t g(t)dt + o(\lambda\tau) = \lambda\tau, \end{aligned} \quad (6.27)$$

где $o(\lambda\tau)$ есть величина малости порядка $(\lambda\tau)^2$.

6.5.2. Показатели надежности

СРЕДНЕЕ ВРЕМЯ РАБОТЫ ДО ОТКАЗА. В соответствии с теоремой разрежения случайного точечного потока среднее время работы до отказа дублированной системы с *холодным* резервом приближенно равняется

$$T_{\text{сист}} \approx \frac{T}{q}, \quad (6.28)$$

где q определяется по (6.27).

Заметим, что (6.28) получено без каких-либо ограничений на вид распределения времени работы до отказа. Если распределение времени работы принять экспоненциальным с параметром λ и положить $\lambda\tau \ll 1$, то (6.28) можно записать в виде

$$T_{\text{сист}} = \frac{T}{\lambda\tau} = \frac{1}{\lambda^2\tau}. \quad (6.29)$$

ВЕРОЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ. Как следует из теоремы о разрежении точечного процесса, результирующий поток отказов системы при малых q становится близким к пуассоновскому, а это значит, что ВБР может быть представлена в виде

$$P(t_0) \approx \exp(-\lambda^2\tau t_0). \quad (6.30)$$

СРЕДНЕЕ ВРЕМЯ ПРОСТОЯ СИСТЕМЫ. Время простоя рассматриваемой дублированной системы зависит от вида распределения времени восстановления и от того, как осуществляется ремонт. В дальнейшем мы будем предполагать, что имеются две ремонтные бригады, способные ремонтировать одновременно оба отказавших элемента в случае необходимости. Рассмотрим два случая:

- ✓ случайное время восстановления элемента подчиняется экспоненциальному распределению со средним τ ;
- ✓ время восстановления элемента постоянно и равно τ .

В любом случае в момент отказа системы ранее отказавший элемент находится в процессе ремонта, а вновь отказавший только начинает восстанавливаться. Если обозначить случайное время ремонта элемента через η , а остаточное время ремонта для того элемента, который на момент отказа системы ремонтировался, через ν , то среднее время простоя системы определяется следующим образом:

$$\tau_{\text{сист}} = \min\{\xi, \nu\}. \quad (6.31)$$

Нетрудно убедиться, что в обоих случаях среднее время восстановления системы равно

$$\tau_{\text{сист}} = \tau/2. \quad (6.32)$$

КОЭФФИЦИЕНТ ГОТОВНОСТИ в общем виде может быть представлен в виде

$$K = \frac{T_{\text{сист}}}{T_{\text{сист}} + \tau_{\text{сист}}}, \quad (6.33)$$

где $T_{\text{сист}}$ и $\tau_{\text{сист}}$ определяются по (6.29) и (6.32). В обоих случаях КГ приближенно равен

$$K \approx 1 - \frac{\tau_{\text{сист}}}{T_{\text{сист}}}. \quad (6.34)$$

В случае экспоненциального распределения времени работы до отказа последнее выражение можно переписать в виде

$$K \approx 1 - \frac{\lambda^2 \tau^2}{2}. \quad (6.35)$$

КОЭФФИЦИЕНТ ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ с использованием (6.30) и (6.35) можно приближенно записать как

$$R(t_0) \approx (1 - 0.5 \lambda^2 \tau^2) \cdot (1 - \lambda^2 \tau t_0) \approx 1 - \lambda^2 \tau (0.5 \tau + t_0). \quad (6.36)$$

ПРИМЕР 6.2. Дублированная система с *холодным* резервом состоит из двух подсистем. Каждая из подсистем представляет собой систему, рассмотренную в примере 6.1. Исходные данные для данного примера, являющиеся результатами решения примера 6.1, сведены в табл. 6.2. Длительность заданной операции системы равна $t_0 = 2$ ч.

В случае необходимости обе подсистемы могут восстанавливаться независимо. Требуется найти основные показатели надежности рассматриваемой дублированной системы.

Решение. Среднее время безотказной работы системы:

$$T_{\text{сист}} \approx [(4.75 \cdot 10^{-3})^2 \cdot 0.737]^{-1} = 60.137 \text{ [ч]}.$$

Интенсивность отказов системы:

$$\lambda_{\text{сист}} = 60.137^{-1} = 1.66 \cdot 10^{-5} \text{ [1/ч]}.$$

Среднее время простоя системы:

$$\tau_{\text{сист}} = 0.737/2 = 0.369 \text{ [ч]}.$$

Вероятность безотказной работы для операции длительностью 2 ч:

$$P(2) \approx 1 - 2 \cdot 1.66 \cdot 10^{-5} \approx 0.999967.$$

Таблица 6.2

Среднее время безотказной работы T , [ч]	210.5
Интенсивность отказов λ , [1/ч]	$4.75 \cdot 10^{-3}$
Среднее время восстановления τ , [ч]	0.737

Коэффициент готовности:

$$K \approx 1 - (1.66 \cdot 10^{-5}) \cdot 0.369 = 0.9999939.$$

Коэффициент интервальной готовности для операции длительностью 2 ч:

$$R(2) \approx 0.999967 \cdot 0.9999939 \approx 0.999961.$$

§ 6.6. ВЫСОКОНАДЕЖНОЕ ГОРЯЧЕЕ ДУБЛИРОВАНИЕ

§ 6.6.1. Разрежение потока отказов

Рассмотрим дублированную систему с *горячим* резервом. Оба элемента системы идентичны и характеризуются интенсивностью отказов λ и средним временем восстановления τ . Как только возникает отказ любого элемента, мгновенно начинается восстановление его работоспособности. Переключение с основного элемента на резервный осуществляется также мгновенно. Предполагается, что оба элемента работают независимо, каждый образует свой поток *импульсов* простоя, а отказу системы соответствует совпадение этих импульсов (рис. 6.5). Система отказывает, когда оба элемента оказываются в неработоспособном состоянии.

Определим интенсивность отказов такой дублированной системы, рассмотрев ее коэффициент готовности. Стационарный КГ любой системы может быть выражен через известные значения интенсивности отказов $\lambda_{\text{сист}}$ и среднего времени простоя $\tau_{\text{сист}}$:

$$K_{\text{сист}} = \frac{1}{1 + \lambda_{\text{сист}} \tau_{\text{сист}}} \approx 1 - \lambda_{\text{сист}} \tau_{\text{сист}}. \quad (6.37)$$

В то же время для той же системы, поскольку она состоит из независимых элементов, КГ можно представить и через КГ элементов:

$$K_{\text{сист}} = 1 - (1 - K)^2 \approx 1 - \lambda^2 \tau^2. \quad (6.38)$$

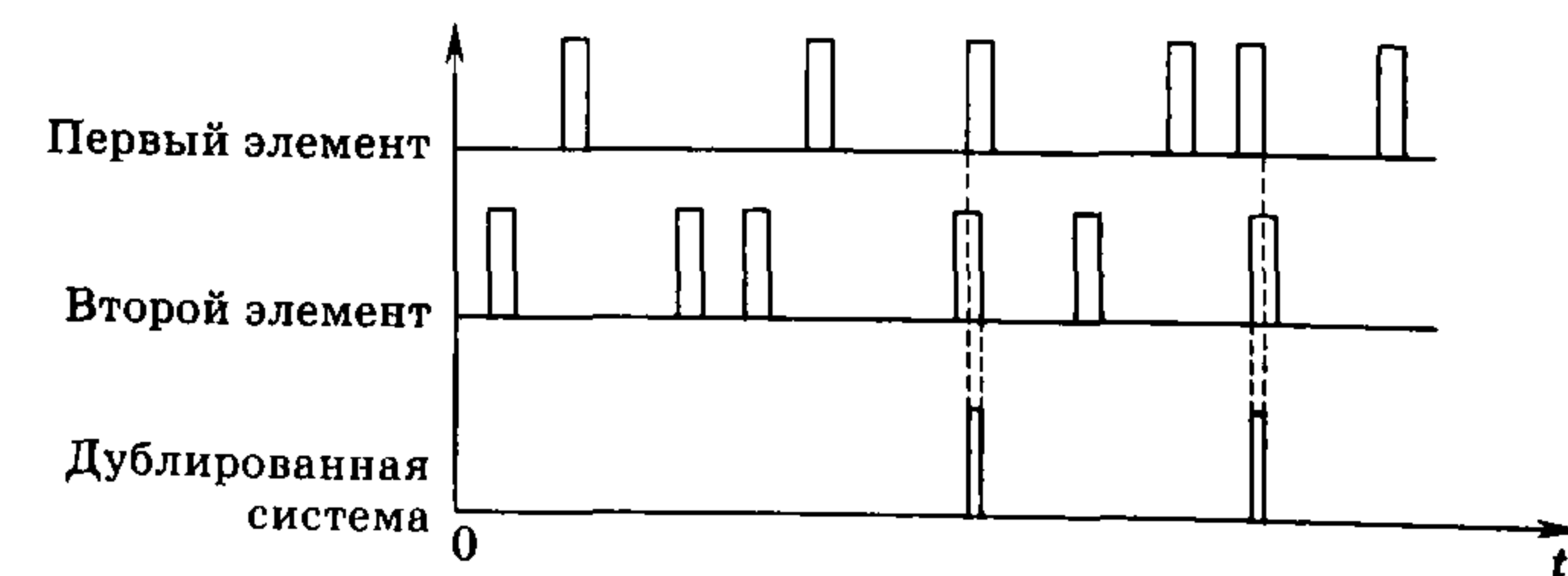


Рис. 6.5

Из (6.37) и (6.38) легко выводится, что

$$\lambda_{\text{сист}} \approx \frac{\lambda^2 \tau^2}{\tau_{\text{сист}}} \quad (6.39)$$

Следовательно, для нахождения $\lambda_{\text{сист}}$ необходимо найти среднее время простоя системы $\tau_{\text{сист}}$. Повторив аргументацию, приведенную в случае с *холодным* дублированием, нетрудно убедиться, что среднее время простоя дублированной системы с *горячим* резервом сохраняется тем же, что было и в случае *холодного* резерва. Это приводит после подстановки в (6.39) к окончательной формуле

$$\lambda_{\text{сист}} \approx 2\lambda^2 \tau. \quad (6.40)$$

В соответствии с теоремой разрежения случайных точечных потоков разреженный поток при высокой степени разрежения можно считать приближенно пуассоновским. Это позволяет легко находить различные показатели надежности систем, рассматриваемые в следующем параграфе.

6.6.2. Показатели надежности

Среднее время безотказной работы с использованием (6.40) можно представить следующим образом:

$$T_{\text{сист}} \approx \frac{1}{2\lambda^2 \tau}. \quad (6.41)$$

ВЕРОЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ, поскольку в процессе разрежения поток отказов стремится к пуассоновскому, может быть выражен как

$$P(t_0) \approx \exp(-2\lambda^2 \tau t_0). \quad (6.42)$$

КОЭФФИЦИЕНТ ГОТОВНОСТИ. Так как было получено, что среднее время простоя системы равно $\tau/2$, а интенсивность отказов определена через (6.40), КГ дублированной системы с *горячим* резервом приближенно равен

$$K \approx \frac{1}{1 + \lambda^2 \tau^2} \approx 1 - \lambda^2 \tau^2. \quad (6.43)$$

КОЭФФИЦИЕНТ ИНТЕРВАЛЬНОЙ ГОТОВНОСТИ в соответствии с (6.42) для малых t_0 может быть приближенно выражен как

$$R(t_0) \approx (1 - \lambda^2 \tau t_0)(1 - \lambda^2 \tau^2) \approx 1 - \lambda^2 \tau (t_0 + \tau). \quad (6.44)$$

Пример 6.3 Дублированная система с *горячим* резервом состоит из двух подсистем, каждая из которых представляет собой систему, рассмотренную в

Таблица 6.3

Среднее время безотказной работы T , ч	210.5
Интенсивность отказов λ , 1/ч	$4.75 \cdot 10^{-3}$
Среднее время восстановления τ , ч	0.737

примере 6.1. Таким образом, исходные данные являются результатами решения примера 6.1. Они сведены в табл. 6.3. Длительность заданной операции системы t_0 равна 2 ч. Обе подсистемы восстанавливаются независимо. Требуется найти основные показатели надежности рассматриваемой дублированной системы с *горячим* резервом.

Решение. Среднее время безотказной работы системы:

$$T_{\text{сист}} \approx [2 \cdot (4.75 \cdot 10^{-3})^2 \cdot 0.737]^{-1} = 30.1 \text{ [ч]}.$$

Интенсивность отказов системы:

$$\lambda_{\text{сист}} = 60.137^{-1} = 3.32 \cdot 10^{-5} \text{ [1/ч]}.$$

Среднее время простоя системы:

$$\tau_{\text{сист}} = 0.737/2 = 0.369 \text{ [ч]}.$$

Вероятность безотказной работы для операции длительностью 2 ч:

$$P(2) \approx 1 - 2 \cdot 3.32 \cdot 10^{-5} \approx 0.999934.$$

Коэффициент готовности:

$$K \approx 1 - (3.32 \cdot 10^{-5}) \cdot (0.369) \approx 0.999988.$$

Коэффициент интервальной готовности для операции длительностью 2 ч:

$$R(2) \approx 0.999934 \cdot 0.999988 \approx 0.999922.$$

§ 6.7. ЭЛЕМЕНТ С ПЕРИОДИЧЕСКИМ КОНТРОЛЕМ РАБОТОСПОСОБНОСТИ

Рассмотрим простейшую систему, состоящую из единственного элемента с произвольным распределением времени работы до отказа, имеющим среднее, равное T . Элемент не имеет постоянного контроля работоспособности, а его состояние проверяется с периодичностью θ и занимает время δ , в течение которого система пребывает в нерабочем состоянии. После проверки система не обновляется, а продолжает функционирование с текущими характеристиками надежности. В этом случае после каждого интервала θ система всегда простаивает в течение времени δ вне зависимости

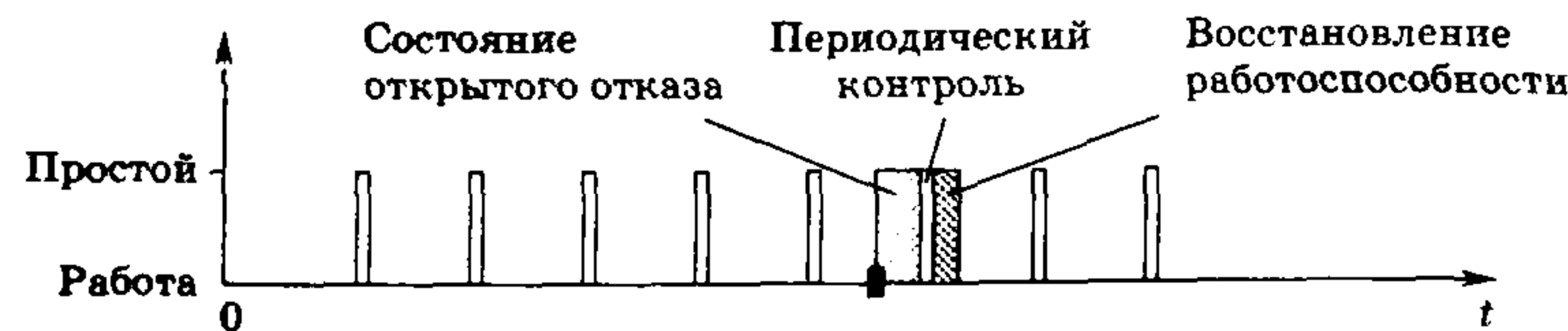


Рис. 6.6

от того, отказала система в течение предыдущего интервала θ или нет (рис. 6.6).

После каждого отказа всегда наступает интервал скрытого состояния работоспособности, который в среднем длится $\theta/2$. Затем осуществляется восстановление работоспособности элемента в среднем в течение времени τ , когда система также простаивает. Составляющая K_1 коэффициента готовности за счет этих факторов приближенно равен

$$K_1(\theta) \approx \frac{T}{T + \theta/2 + \tau} \approx 1 - \frac{\theta + 2\tau}{2T}. \quad (6.45)$$

Кроме того, на каждом цикле контроля коэффициент использования системы K_2 за счет простоев во время проверок равен

$$K_2(\theta) = \frac{\theta}{\theta + \delta} \approx 1 - \frac{\delta}{\theta}. \quad (6.46)$$

Результирующий КГ для случая высокой надежности можно теперь приближенно выразить в виде произведения K_1 на K_2 :

$$K_{\text{сист}}(\theta) \approx \left(1 - \frac{\theta + 2\tau}{2T}\right) \left(1 - \frac{\delta}{\theta}\right) \approx 1 - \left(\frac{\theta + 2\tau}{2T} + \frac{\delta}{\theta}\right). \quad (6.47)$$

ПРИМЕР 6.4. Длительность каждой проверки составляет $\delta = 0.2$ ч, среднее время безотказной работы системы $T = 100$ ч, время восстановления $\tau = 2$ ч. Требуется найти периодичность проверок θ , которая обеспечивала бы максимальное значение КГ.

Решение. Используя формулу (6.47), проведем вычисления и сведем их в таблицу. Как следует из таблицы, при заданных параметрах оптимальным является проведение проверок через 6 ч (максимальное значение КГ выделено жирным шрифтом).

Таблица 6.4

θ	1	2	3	4	5	6	7	8	9
КГ	0.775	0.87	0.8983	0.91	0.915	0.9167	0.9164	0.915	0.9128

§ 6.8. УЧЕТ ВРЕМЕНИ ПЕРЕКЛЮЧЕНИЯ ПРИ ДУБЛИРОВАНИИ С ХОЛОДНЫМ РЕЗЕРВОМ

Рассмотрим дублированную систему с ненагруженным (*холодным*) резервом, у которой подключение резервного элемента вместо отказавшего основного занимает время δ . Оба элемента могут восстанавливаться одновременно. Среднее время восстановления равно τ . В этом случае поток отказов системы будет происходить:

- ✓ каждый раз после отказа основного элемента, т. е. в среднем через $T = 1/\lambda$ ч система простаивает в течение интервала δ из-за переключения на резерв;
- ✓ каждый раз после отказа обоих элементов одновременно система простаивает в среднем в течение времени $\tau/2$, пока не будет восстановлен один из отказавших элементов.

Рассматривая высоконадежную систему, мы пренебрегаем возможностью перекрытия во времени обоих событий.

Для данного случая можно составить условную блок-схему надежности, представленную на рис. 6.7.

Таким образом, мы имеем последовательное соединение идеальной дублированной системы с *холодным* резервом и условного элемента (переключателя). Дублированная система характеризуется коэффициентом готовности

$$K_1 \approx 1 - 0.5 \cdot \lambda^2 \tau^2. \quad (6.48)$$

Переключатель отказывает с интенсивностью основного элемента и характеризуется временем простоя δ , т. е. его КГ, K_2 , приближенно равен

$$K_2 \approx 1 - \lambda \delta. \quad (6.49)$$

Это позволяет записать приближенное выражение для КГ такой системы в виде

$$K_{\text{сист}} \approx K_1 K_2 \approx (1 - 0.5 \cdot \lambda^2 \tau^2) (1 - \lambda \delta) \approx 1 - \lambda(0.5 \cdot \lambda \tau^2 + \delta). \quad (6.50)$$

ПРИМЕР 6.5. Элемент дублированной системы с *холодным* резервом имеет $T = 1000$ ч и $\tau = 1$ ч. Переключение на резерв занимает время $\delta = 0.01$ ч. Требуется найти КГ системы.

Решение. Коэффициент готовности составляет $K \approx 1 - 0.001(0.5 \cdot 0.001 \cdot 1 + 0.01) + 0.01 = 0.9999895$. Заметим, что КГ дублированной системы без учета времени переключения равен

$K \approx 1 - 0.5 \cdot 0.0001 \cdot 0.0001 = 0.9999995$, т. е. коэффициент простоя, полученный по обычной формуле для *холодного* дублирования, занижен по сравнению с точной формулой почти в 20 раз.

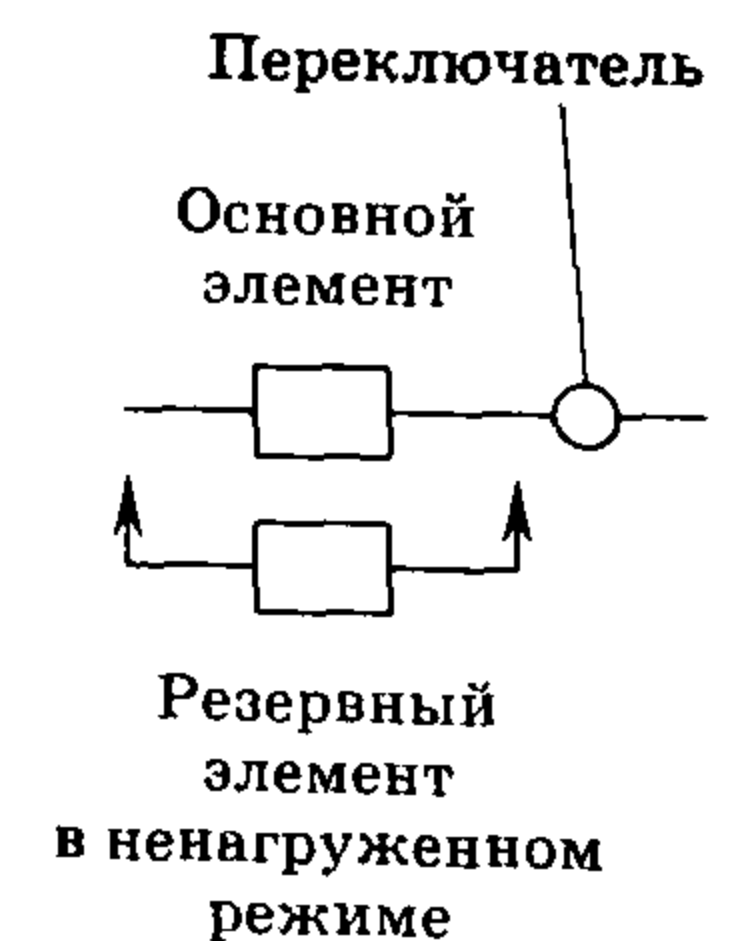


Рис. 6.7

§ 6.9. УЧЕТ ВРЕМЕНИ ПЕРЕКЛЮЧЕНИЯ ПРИ ДУБЛИРОВАНИИ С ГОРЯЧИМ РЕЗЕРВОМ

Рассмотрим дублированную систему с нагруженным (*горячим*) резервом, у которой переключение с отказавшего основного элемента на резервный занимает время δ . Опять возможно одновременное восстановление обоих отказавших элементов. Среднее время восстановления отказавшего элемента равно τ . В этом случае поток отказов системы будет происходить:

- ✓ каждый раз после отказа основного элемента, т. е. в среднем через $T = 1/\lambda$ ч, система простаивает в течение интервала δ из-за переключения на резерв;
- ✓ каждый раз после отказа обоих элементов одновременно система простаивает в среднем в течение времени $\tau/2$, пока не восстановится один из ее элементов.

Рассматривая высоконадежную систему, мы пренебрегаем возможностью перекрытия во времени этих двух событий.

Для данного случая можно составить условную блок-схему надежности, представленную на рис. 6.8. Здесь мы имеем последовательное соединение идеальной дублированной системы с *горячим* резервом и условного элемента (переключателя). КГ дублированной системы приближенно равен

$$K_1 \approx 1 - \lambda^2 \tau^2. \quad (6.51)$$

Переключатель отказывает с интенсивностью основного элемента и характеризуется временем простоя δ , т. е. его КГ приближенно равен

$$K_2 \approx 1 - \lambda \delta. \quad (6.52)$$

Это позволяет получить приближенное выражение для КГ дублированной системы K в виде

$$K \approx K_1 K_2 \approx (1 - \lambda^2 \tau^2) (1 - \lambda \tau) \approx 1 - \lambda(\lambda \tau^2 + \delta). \quad (6.53)$$

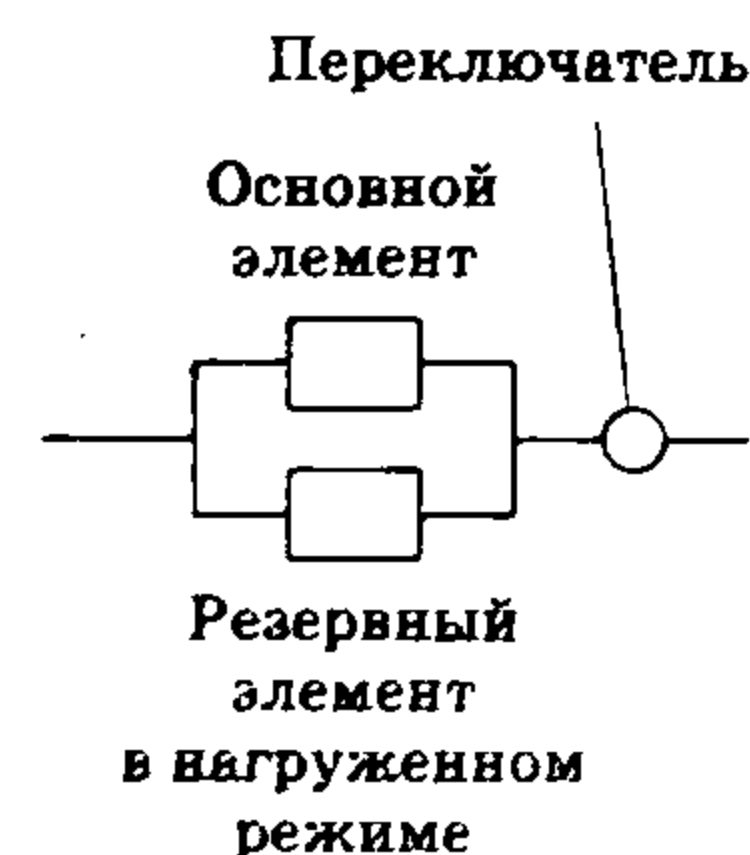


Рис. 6.8

ПРИМЕР 6.6. Элемент дублированной системы с *горячим* резервом имеет следующие характеристики надежности: $T = 1000$ ч и $\tau = 1$ ч. Переключение на резерв занимает время $\delta = 0.01$ ч. Требуется найти КГ системы.

Решение. Коэффициент готовности составляет

$$K \approx 1 - 0.001(0.001 \cdot 1 + 0.01) = 0.999989.$$

Заметим, что КГ той же дублированной системы без учета времени переключения составил бы

$$K \approx 1 - 0.001 \cdot 0.001 = 0.999999,$$

т. е. коэффициент простоя занижен по отношению к реальному более чем в 10 раз.

§ 6.10. НЕПОЛНЫЙ КОНТРОЛЬ ОСНОВНОГО ЭЛЕМЕНТА ДУБЛИРОВАННОЙ СИСТЕМЫ С ГОРЯЧИМ РЕЗЕРВОМ

Рассмотрим дублированную систему с восстановлением, у которой основной элемент непрерывно контролируется только на $\alpha \cdot 100\%$, $\alpha < 1$, а резервный элемент, находящийся в *горячем* режиме, контролируется полностью и непрерывно. Говоря о неполном контроле, мы имеем в виду, что *элементом* может быть, как уже отмечалось, достаточно сложная часть аппаратуры, включающая в свой состав большое число «атомарных» компонентов. Условная блок-схема расчета надежности такой дублированной системы представлена на рис. 6.9.

Периодически через интервал θ производится проверка основного элемента полностью. Это можно осуществить, в частности, переключаясь на резервный элемент, состояние которого известно всегда. Сами по себе такие проверки нельзя проводить слишком часто, потому что на каждую такую проверку затрачивается конечное (хотя, возможно, и очень небольшое) время, а кроме того, всякая проверка нарушает выполнение системой текущей операции.

Пусть элементы сами по себе идентичны. Случайное время работы до отказа имеет экспоненциальное распределение с параметром λ , а среднее время восстановления равно τ . Предполагается, что $\lambda \tau \ll 1$. Рассматриваемая система может находиться в состоянии простоя в двух следующих случаях:

- (1) на интервале θ отказывает неконтролируемая часть основного элемента, и система оказывается в состоянии скрытого отказа;
- (2) на интервале θ одновременно отказываются отказавшими и контролируемая часть основного элемента, и резервный элемент.

В случае (1) интенсивность потока отказов неконтролируемой части основного элемента равна $(1 - \alpha)\lambda$. Если среднее время безотказной рабо-

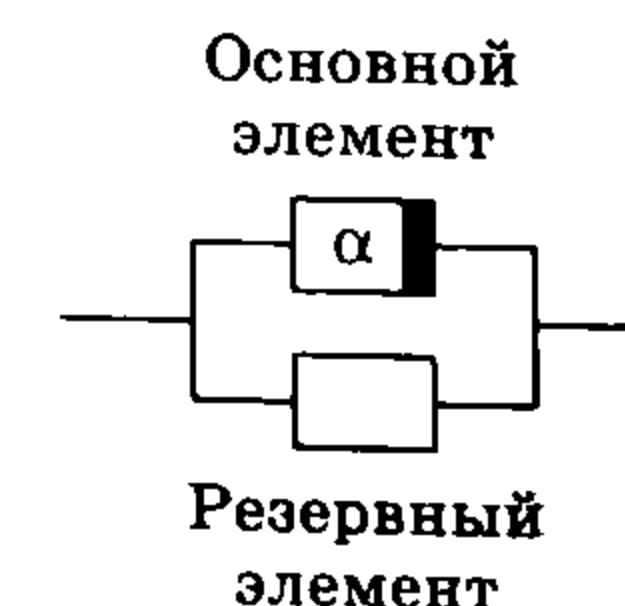


Рис. 6.9

ты элемента много больше, чем интервал периодических проверок θ , то можно считать, что момент отказа неконтролируемой части основного элемента распределен равномерно внутри интервала θ . Это означает, что длительность простоя системы с необнаруженным отказом равна в среднем 0.5θ .

Составляющая K_1 коэффициента готовности системы за счет этого фактора равна

$$K_1 \approx 1 - (1 - \alpha)\lambda \cdot \theta / 2. \quad (6.54)$$

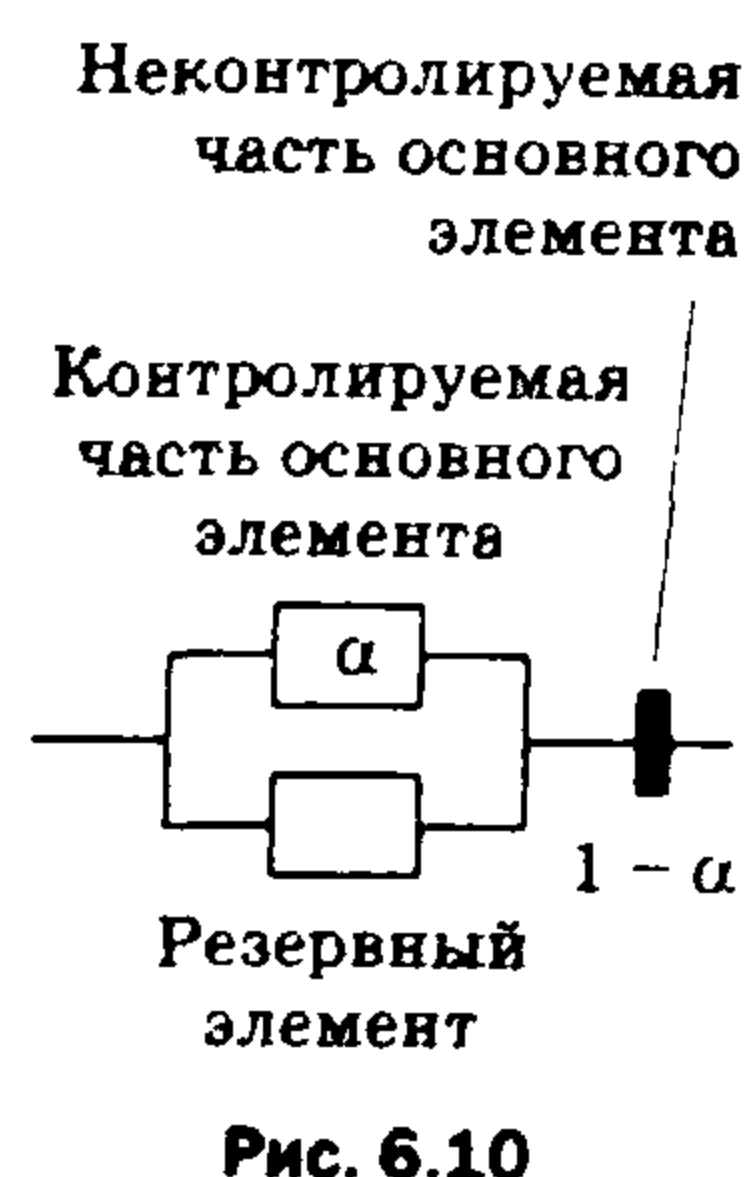
В случае (2) можно записать для двух независимых элементов выражение

$$K_2 \approx 1 - (\lambda\tau)(\alpha\lambda\tau) = 1 - \alpha\lambda^2\tau^2. \quad (6.55)$$

Приближенная итоговая формула КГ системы имеет вид

$$K \approx K_1 K_2 \approx 1 - [0.5(1 - \alpha)\lambda\theta + \alpha\lambda^2\tau^2]. \quad (6.56)$$

Применительно к этой приближенной модели можно предложить блок-схему надежности (рис. 6.10), которая грубо описывает эту ситуацию*. Здесь в режиме обычного *горячего* резерва включены параллельно два элемента и последовательно им включен элемент, который отказывает с интенсивностью $(1 - \alpha)\lambda$ и простаивает в состоянии необнаруженного отказа в среднем в течение интервала времени $\theta/2$. Стационарный коэффициент готовности, полученный для такой эквивалентной блок-схемы надежности, естественно, совпадает с представленными в (6.56), поскольку допущения о высокой надежности в обоих случаях совпадают.



ПРИМЕР 6.7. Рассмотрим дублированную систему с неполным контролем основного элемента. Пусть $\lambda = 0.001$ 1/ч; $\alpha = 95\%$; $\tau = 0.5$ ч; $\theta = 20$ ч. Требуется найти КГ системы.

Решение. При заданных значениях параметров КГ системы равен

$$K \approx 1 - (0.5 \cdot 0.05 \cdot 10^{-3} \cdot 20 + 0.95 \cdot 10^{-6} \cdot 0.25) \approx 0.9995.$$

Заметим, что расчет при нагруженном дублировании без учета неполноты контроля дает сильное завышение:

$$K \approx 1 - 0.25 \cdot 10^{-6} = 0.99999975.$$

* Одновременное осуществление обоих рассмотренных выше событий не учитывается, поскольку для высоконадежной системы это крайне маловероятно.

Сравнивая вычисленные значения коэффициентов простоя системы для двух приведенных схем расчета и вводя коэффициент k , $k = 1 - K$, видим, что расчет идеализированной модели ($k = 0.25 \cdot 10^{-6}$) по сравнению с уточненной моделью ($k = 5 \cdot 10^{-4}$) дает двадцатикратную ошибку!

§ 6.11. НЕПОЛНЫЙ КОНТРОЛЬ РЕЗЕРВНОГО ЭЛЕМЕНТА ДУБЛИРОВАННОЙ СИСТЕМЫ С ГОРЯЧИМ РЕЗЕРВОМ

Рассмотрим дублированную систему с восстановлением, у которой основной элемент контролируется полностью и непрерывно, а резервный элемент, находящийся в *горячем* режиме, непрерывно контролируется только на $\alpha \cdot 100\%$, $\alpha < 1$. Говоря о неполном контроле *элемента*, мы опять подразумеваем, что элементом может быть некая подсистема с достаточно сложной структурой и большим числом *атомарных* компонентов.

Условная блок-схема расчета надежности такой дублированной системы представлена на рис. 6.11.

Периодически через время θ производится проверка резервного элемента полностью.

Пусть каждый из элементов имеет экспоненциальное распределение времени работы до отказа с параметром λ и среднее время восстановления τ . Рассматриваемая система может отказаться в двух случаях:

- (1) на интервале θ отказывают и основной элемент, и контролируемая часть резервного элемента, причем интервалы их восстановлений пересекаются;
- (2) на интервале θ отказывает основной элемент в момент, когда неконтролируемая часть резервного элемента уже отказала, т. е. резервный элемент находится в состоянии скрытого отказа. (Если отказ основного элемента происходит раньше, то при условии $\theta \gg \tau$ можно считать, что отказ успевает восстановиться.)

Случай (1) может быть описан обычной дублированной системой, основной и резервный элементы которой имеют интенсивности отказов соответственно λ и $\alpha\lambda$. Этот случай совпадает с представленным формулой (6.55), т. е. КГ системы может быть записан опять в виде

$$K_1 \approx 1 - \alpha\lambda^2\tau^2. \quad (6.57)$$

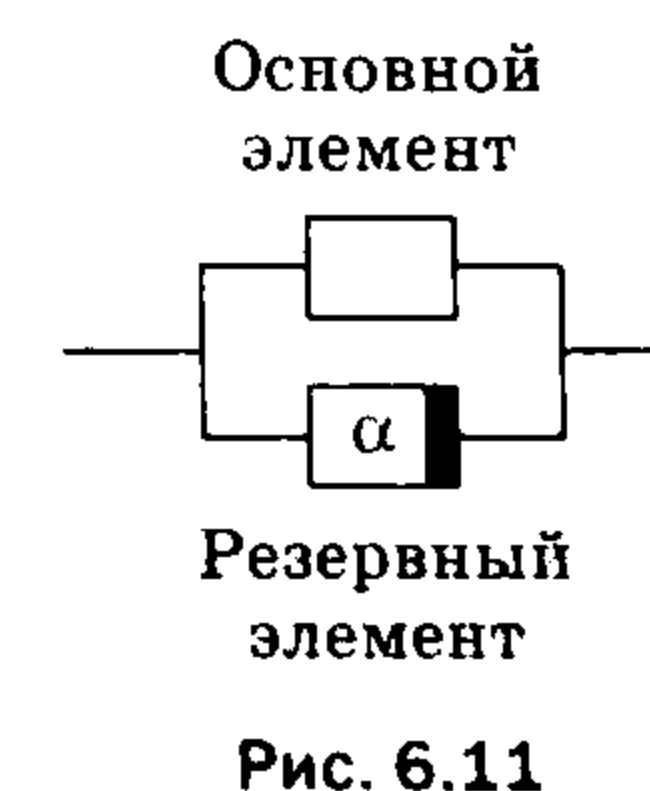


Таблица 6.5

Число ремонтных органов	Вид распределения времени восстановления	
	Экспоненциальное	Вырожденное
1	τ	τ
2	0.5τ	τ

В случае (2) интенсивность потока отказов неконтролируемой части резервного элемента равна $(1 - \alpha)\lambda$, а вероятность того, что в течение скрытого отказа резервного элемента произойдет отказ основного элемента, приближенно равняется $0.5\lambda\theta$. Иначе говоря, интенсивность отказов, вызванных отказом основного элемента и неконтролируемой части резервного элемента, приближенно равна $(1 - \alpha)\lambda \cdot (0.5 \cdot \lambda\theta)$.

Среднее время простоя системы в этом случае зависит от числа ремонтных органов и от вида распределения времени восстановления. Результирующие значения среднего времени восстановления сведены в табл. 6.5.

Пояснение требуется, возможно, лишь для вырожденного распределения времени восстановления (время восстановления постоянно): в этом случае время простоя системы не зависит от числа ремонтных органов. Будем полагать в дальнейшем для простоты, что время простоя системы в данной ситуации равно τ . Тогда составляющая КГ, определяемая неконтролируемой частью резервного элемента, может быть представлена в виде

$$K_2 = 1 - 0.5(1 - \alpha)\lambda^2\tau\theta. \quad (6.58)$$

В результате КГ системы равен

$$K = K_1K_2 \approx 1 - \lambda^2\tau[\alpha\tau + 0.5 \cdot (1 - \alpha)\theta]. \quad (6.59)$$

ПРИМЕР 6.8. Рассмотрим дублированную систему с неполным контролем резервного элемента. Пусть $\lambda = 0.001$ 1/ч; $\alpha = 95\%$; $\tau = 1$ ч; $\theta = 20$ ч. Требуется найти КГ системы.

При заданных значениях параметров КГ системы равен

$$K \approx 1 - 0.000001(0.95 + 0.5 \cdot 0.05 \cdot 20) \approx 0.99999855.$$

Заметим, что расчет «классического» нагруженного резервирования дает $K = 1 - 1 \cdot 10^{-6} = 0.999999$.

Сравнивая вычисленные значения коэффициентов простоя системы для двух приведенных схем расчета, видим, что они разнятся почти в два раза.

§ 6.12. УЧЕТ НЕНАДЕЖНОСТИ ПЕРЕКЛЮЧЕНИЯ

Для конкретности рассмотрим дублированную систему с *горячим* резервом и единственным ремонтным органом. Время восстановления работоспособности отказавшего элемента равно τ . Переключение с отказавшего основного элемента на резервный происходит успешно с вероятностью p , а с вероятностью $q = 1 - p$ переключение не срабатывает, и требуется время π^* для восстановления переключающего устройства. Переключения предполагаются независимыми, а вероятности отказов при переключениях постоянными.

В этом случае возможны две следующие ситуации:

- ✓ при условии, что переключатель идеален, рассматривается обычная дублированная система с *горячим* резервом, КГ которой

$$K_1 \approx 1 - \lambda^2\tau^2; \quad (6.60)$$

- ✓ после каждого отказа основного элемента, переключатель отказывает с вероятностью q и требует для своего восстановления времени π , что характеризуется составляющей КГ

$$K_2 \approx 1 - \lambda q\pi. \quad (6.61)$$

Если и на этот раз пренебречь малой вероятностью того, что указанные события могут пересечься во времени, то результирующий КГ системы приближенно запишется как

$$K = K_1K_2 \approx 1 - \lambda(\lambda\tau^2 + q\pi). \quad (6.62)$$

ПРИМЕР 6.9. Рассмотрим дублированную систему с ненадежным переключателем. Пусть $\lambda = 0.001$ 1/ч; $\tau = 1$ ч; вероятность успешного срабатывания переключателя $p = 0.99$, а время его восстановления $\pi = 0.5$ ч. Требуется найти КГ системы.

Решение. При заданных значениях параметров КГ системы равен

$$K \approx 1 - 0.001 \cdot (0.001 + 0.01 \cdot 0.5) \approx 0.999994.$$

Заметим, что расчет для дублирования с *горячим* резервом без учета переключателей дает

$$K = 1 - 1 \cdot 10^{-6} = 0.999999.$$

Таким образом, неучет отказов при переключениях приводит к неоправданно завышенным результатам.

Применение эвристического метода анализа восстанавливаемых систем не ограничивается приведенными примерами. Этот метод успешно применяется и в гораздо более сложных моделях систем с восстановлением.

* Здесь и далее буква π никак не связана с числом 3.1415... — *Ред.*

Глава 7

Оценка эффективности функционирования сложных систем

Для систем с несколькими уровнями функционирования обычный подход оценки надежности оказывается недостаточным — не удается сформулировать критерий отказа, так как по мере отказов элементов качество функционирования системы лишь постепенно деградирует. Предлагается метод анализа эффективности функционирования подобных систем.

§ 7.1. НЕОБХОДИМОСТЬ РАСШИРЕНИЯ ПОНЯТИЯ НАДЕЖНОСТИ

Современные большие системы с их сложной структурой, внутренней избыточностью и многофункциональностью далеко не всегда могут быть охарактеризованы рассмотренными выше «чистыми» показателями надежности. Отказы отдельных элементов системы не обязательно приводят к отказу системы, а могут лишь понизить эффективность функционирования: система может продолжать работать, например, с более низкой производительностью или с худшим качеством выполнения своих операций. Такие системы удобно характеризовать коэффициентом эффективности функционирования, который определяет средний уровень качества функционирования системы или выполнения ею своих операций, обусловленный конечной надежностью входящих в систему элементов.

Часто коэффициент эффективности функционирования оказывается размерным, поскольку выходной эффект системы может измеряться в различных единицах: в объеме производимой продукции, в скорости выполнения операции, в доле работоспособ-

ных конечных устройств в системах управления и т. д. Однако почти всегда удобнее ввести безразмерный коэффициент эффективности, который показывает относительное снижение эффективности функционирования системы за счет отказов ее элементов по сравнению с абсолютно надежной системой.

Безусловно, и системы с деградацией эффективности функционирования можно характеризовать такими показателями надежности, как средняя наработка или коэффициент готовности: для этого достаточно ввести конкретный критерий отказа системы. Например, система считается отказавшей, если ее производительность упала на 2%. Можно ввести целый набор таких показателей для нескольких уровней производительности: для 5% или 10% снижения производительности и т. п. Однако, как уже отмечалось в гл. 1 при обсуждении выбора показателей надежности, наличие большого числа показателей не проясняет картину, а только ее запутывает. Поэтому для сложных систем использование коэффициента эффективности функционирования принято наиболее предпочтительным.

§ 7.2. КЛАССЫ СИСТЕМ

Рассмотрим систему, состоящую из n элементов. Обозначим состояние i -го элемента системы через x_i . Как и прежде, каждый элемент системы может находиться в двух состояниях: работоспособности ($x_i = 1$) и отказа ($x_i = 0$). Всего различных состояний системы, естественно, 2^n . Обозначим состояние системы через $X = (x_1, x_2, \dots, x_n)$.

Пусть система функционирует в течение длительного времени, выполняя определенные операции. Принимая во внимание длительность этих операций, все системы удобно подразделить на два класса: системы кратковременного действия и системы длительного действия. В первом случае имеется в виду, что за время операции в системе практически не происходит изменения ее состояния:

$$P\{X(t) = X(t + t_0)\} = 1 - \epsilon, \quad (7.1)$$

где t_0 — длительность операции, выполняемой системой, а $\epsilon \ll 1$.

Для систем кратковременного действия в каждом состоянии достаточно знать мгновенное значение эффективности функционирования системы (ее *выходного эффекта*) в текущем состоянии.

Тогда коэффициент эффективности функционирования определяется как среднее значение выходного эффекта по всем состояниям системы.

Когда условие (7.1) не выполняется, рассматривается поведение системы, т. е. траектория ее переходов из состояния в состояние. В этом случае надо уметь характеризовать количественно всякую траекторию, а потом проводить усреднение по всем траекториям. Оценка эффективности функционирования подобного рода систем весьма трудоемка и сложна и здесь опускается.

§ 7.3. СИСТЕМЫ КРАТКОВРЕМЕННОГО ДЕЙСТВИЯ

Обозначим через $h_{X_k}(t)$ вероятность того, что система кратковременного действия находится в состоянии X_k в момент времени t . Обозначим условную эффективность системы при условии, что она находится в состоянии X_k , через W_{X_k} . Тогда коэффициент эффективности системы в соответствии с записанным выше определением составит

$$W_{\text{сист}}(t) = \sum_{k=1}^{2^n} h_{X_k}(t) W_{X_k}. \quad (7.2)$$

Будем в дальнейшем опускать зависимость показателя эффективности от момента выполнения системой своей операции, тем более что на практике обычно для сложных систем рассматривается стационарный режим.

Если элементы системы независимы, то вероятность для любого состояния X_k составляет

$$h_{X_k} = \prod p_i^{x_i} q_i^{1-x_i}. \quad (7.3)$$

Введем дополнительные обозначения: h_0 — вероятность того, что в системе работоспособны все элементы:

$$h_0 = \prod_{i=1}^n p_i,$$

h_i — вероятность того, что в системе отказал лишь элемент i :

$$h_i = q_i \prod_{j=1, j \neq i}^n p_j,$$

h_{ij} — вероятность того, что в системе отказали два элемента i и j :

$$h_{ij} = q_i q_j \prod_{\substack{k=1 \\ k \neq (i,j)}}^n p_k$$

и т. д. Введем также обозначение $g_i = q_i p_i$. Тогда можно записать

$$W_{\text{сист}} = h_0 \left(W_0 + \sum_{i=1}^n W_i g_i + \sum_{i,j=1}^n W_{ij} g_i g_j + \dots \right). \quad (7.4)$$

При больших значениях n расчет эффективности в общем случае представляет собой сложную, а то и нерешаемую, вычислительную задачу, как всякая задача переборного характера большой размерности. Однако, если $\max q_i \ll 1/n$, можно предложить метод приближенного вычисления эффективности системы:

$$W_{\text{сист}} \approx h_0 \left(W_0 + \sum_{i=1}^n W_i q_i \right), \quad (7.5)$$

где вместо g_i используется непосредственно величина q_i , поскольку величина p_i близка к единице.

Как отмечалось, удобно рассматривать относительный коэффициент эффективности функционирования системы: $\tilde{W}_{\text{сист}} = W_{\text{сист}}/W_0$. В этом случае (7.5) можно переписать как

$$\begin{aligned} \tilde{W}_{\text{сист}} &\approx \left(1 - \sum_{i=1}^n q_i \right) \cdot \left(1 + \sum_{i=1}^n \tilde{W}_i q_i \right) \approx \\ &\approx 1 - \sum_{i=1}^n q_i + \sum_{i=1}^n \tilde{W}_i q_i = 1 - \sum_{i=1}^n \tilde{w}_i q_i, \end{aligned} \quad (7.6)$$

где $\tilde{w}_i = 1 - \tilde{W}_i$.

Величину \tilde{w}_i часто называют коэффициентом важности элемента. В общем случае \tilde{w}_i зависит не только от структурного положения самого элемента, но и от состояния других элементов; он характеризует важность элемента при полностью работоспособном состоянии остальных элементов системы. Если элементы имеют экспоненциальное распределение времени до отказа $p_i(t) = \exp(-\lambda_i t)$, то (7.6) можно переписать:

$$\tilde{W}_{\text{сист}}(t) \approx 1 - \sum_{i=1}^n \tilde{w}_i \lambda_i t \approx \left(- \sum_{i=1}^n \tilde{w}_i \lambda_i t \right). \quad (7.7)$$

В этом случае важность элемента отражается на снижении величины интенсивности его отказа.

Если рассматривается восстанавливаемая система, у которой каждый элемент характеризуется своим коэффициентом готовности $K_i = 1/(1 + \lambda_i \tau_i)$, то приближенную формулу (7.6) можно записать в виде

$$\tilde{W}_{\text{сист}} \approx 1 - \sum_{i=1}^n \lambda_i \tau_i \tilde{w}_i. \quad (7.8)$$

В этом случае можно считать, что $\lambda_i \tilde{w}_i$ представляет собой интенсивность отказов i -го элемента с учетом его структурной важности.

При проведении мероприятий по повышению надежности коэффициент важности используется для выделения тех элементов,

которые наиболее существенно влияют на результирующую эффективность системы.

Лучше всего проиллюстрировать метод расчета коэффициента эффективности с помощью простых конкретных примеров, в которых традиционные расчеты надежности оказываются неприменимыми.

ПРИМЕР 7.1. Система управления воздушным движением в аэропорту включает два радиолокатора, эффективная зона действия каждого из которых составляет 180° (рис. 7.1).

Предположим, что вероятность успешного выполнения задачи радаром в своей зоне действия равна $p = 0.9$, т. е. даже при работоспособности обеих станций система выполняет лишь 90% задач. Пусть КГ каждого радара равен $K = 0.9$. Допустим, что эффективность системы падает до 50% от номинального значения, если откажет один из радаров. Требуется оценить коэффициент эффективности системы.

Решение. Коэффициент эффективности системы составляет

$$W_{\text{сист}} = (0.9)K^2 + 2 \cdot (0.5) \cdot (0.9) \cdot K(1 - K) = \\ = (0.9) \cdot (0.95^2) + (0.9) \cdot (0.95) \cdot (0.05) = 0.855.$$

Если не учитывать того, что отказы в системе приводят лишь к частичной потере эффективности, то надежность системы охарактеризовалась бы величиной $pK = 0.81225$.

Эту задачу можно решать и несколько иным способом. Нетрудно заметить, что система распадается на две независимые части, каждая из которых покрывает половину всех функций системы. Итак, каждая подсистема вносит вклад, равный

$$(0.5) \cdot (0.9) \cdot (0.95) = 0.4275,$$

а следовательно, в целом система обеспечивает коэффициент эффективности, равный $2 \cdot (0.4275) = 0.855$.

ПРИМЕР 7.2. Рассмотрим аналогичную систему управления воздушным движением, но включающую уже четыре радиолокационных станции. Пусть КГ каждого радара $K = 0.95$, $p = 0.95$ и в любой зоне, обслуживаемой двумя

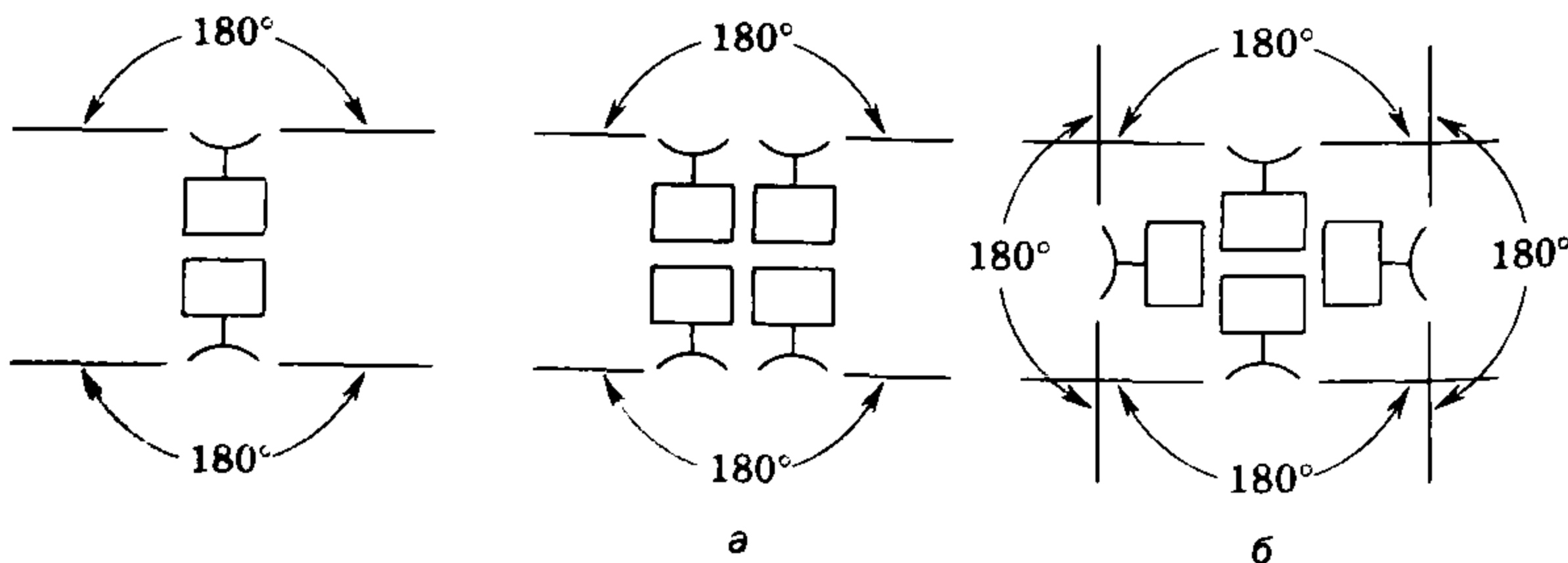


Рис. 7.1

Рис. 7.2

радары, мера эффективности $p^* = 1 - (1 - p)^2 = 0.99$. Имеется две возможных конфигурации такой системы, которые представлены на рис. 7.2. Требуется оценить эффективность обоих вариантов и сравнить их.

Решение. Вариант а представляет собой две независимые подсистемы, каждая из которых вносит свою половину вклада в эффективность системы. В этом случае можно записать формулу полной вероятности для подсчета коэффициента готовности в виде

$$W_{\text{сист}} = [1 - (1 - p)^2]K^2 + p \cdot 2K(1 - K) = \\ = 0.99 \cdot 0.81 + 0.9 \cdot 2 \cdot 0.95 \cdot 0.05 \approx 0.893.$$

Вариант б является более интересным с точки зрения проведения анализа эффективности. Здесь возможны несколько типов ситуаций, представленных на рис. 7.3*. На этом рисунке черный треугольник показывает зону отказавшего радара, темно-серый фон соответствует зонам перекрытия двух радаров (коэффициент эффективности равен 0.99), светло-серый фон — зоне действия единственного радара (условный коэффициент эффективности равен 0.9) и незаштрихованная зона — отсутствию радиолокационного обслуживания в зоне.

Результирующий коэффициент эффективности системы равен

$$W_{\text{сист}} = R_a W_a + R_b W_b + R_c W_c + R_d W_d + R_e W_e = \\ = K^4 \cdot [1 - (1 - p)^2] + 4K^3(1 - K) \{0.5p + 0.5 \cdot [1 - (1 - p)^2]\} + \\ + 4K^2(1 - K)^2 \{0.25 \cdot [1 - (1 - p)^2] + 0.75p\} + \\ + 2K^2(1 - K)^2 p \cdot 4K(1 - K)^3 \cdot (0.5p) \approx 0.957.$$

Таким образом, преимущество второго варианта бесспорно: вероятность невыполнения системой своих функций снизилась с 0.107 до 0.043, т. е. более чем в пять раз. В таких случаях в обиходе часто говорят, что «надежность повысилась в пять раз»: нужно понимать, что речь идет о соответствующем снижении, а не о повышении надежности.

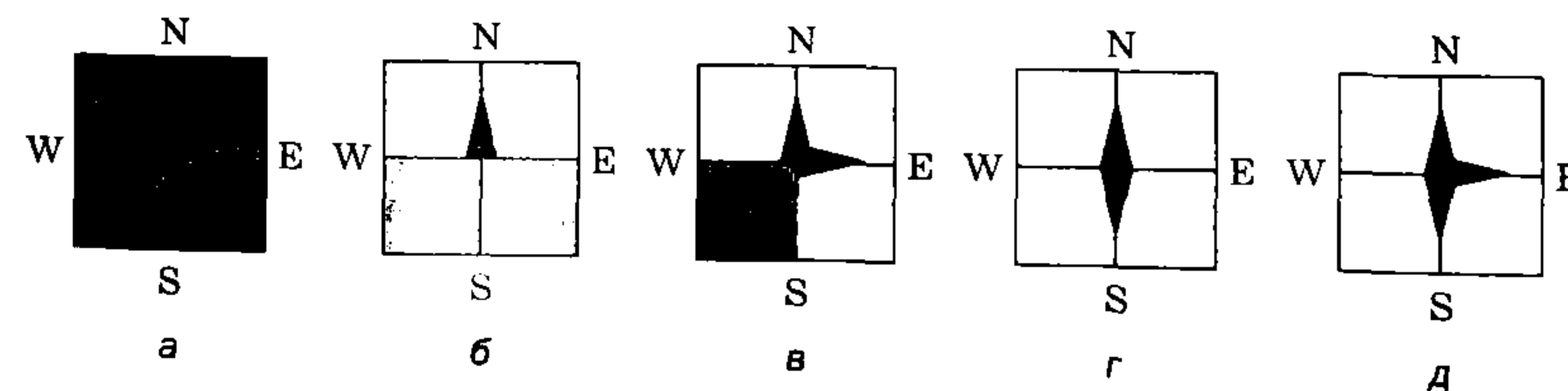


Рис. 7.3

* На рис. 7.3 показано по одному варианту выбора единственного отказавшего радара, двух соседних отказавших радаров, двух радаров, расположенных «спина к спине», и, наконец, трех отказавших радаров. В действительности один радар можно выбрать четырьмя способами, два соседних — четырьмя, три радара — тоже четырьмя, и, наконец, два, расположенных «спина к спине», — двумя способами.

§ 7.4. СИСТЕМЫ С ВЕТВЯЩЕЙСЯ СТРУКТУРОЙ

Многие системы имеют ветвящуюся структуру: они состоят из исполнительных элементов, которые управляются элементами высшего уровня иерархии управления; те, в свою очередь, элементами следующего уровня управления и т. д. Такую структуру имеют большие системы вооружения (практически все командные системы, начиная от управления сухопутными войсками и кончая управлением безлюдными пусковыми установками баллистических ракет), энергетические системы, транспортные системы и др. Пример ветвящейся структуры приведен на рис. 7.4. На этом рисунке элементы с 1 по 6 являются управляющими, а элементы с 7 по 15 — исполнительными. Отказ управляющего элемента приводит к отказу всех подчиненных ему исполнительных элементов, даже если они сами по себе и работоспособны. Например, отказ элемента 2 приводит к отказам исполнительных элементов с 7 по 11, а отказ элемента 1 приводит к отказу всех элементов системы. Исполнительные элементы оказываются зависимыми через общие управляющие элементы, поэтому оценка коэффициента эффективности таких систем оказывается достаточно нетривиальной.

7.4.1. Система с коэффициентом эффективности аддитивного типа

Рассмотрим систему с ветвящейся структурой, каждый исполнительный элемент которой вносит свою независимую долю в общий выходной эффект системы.

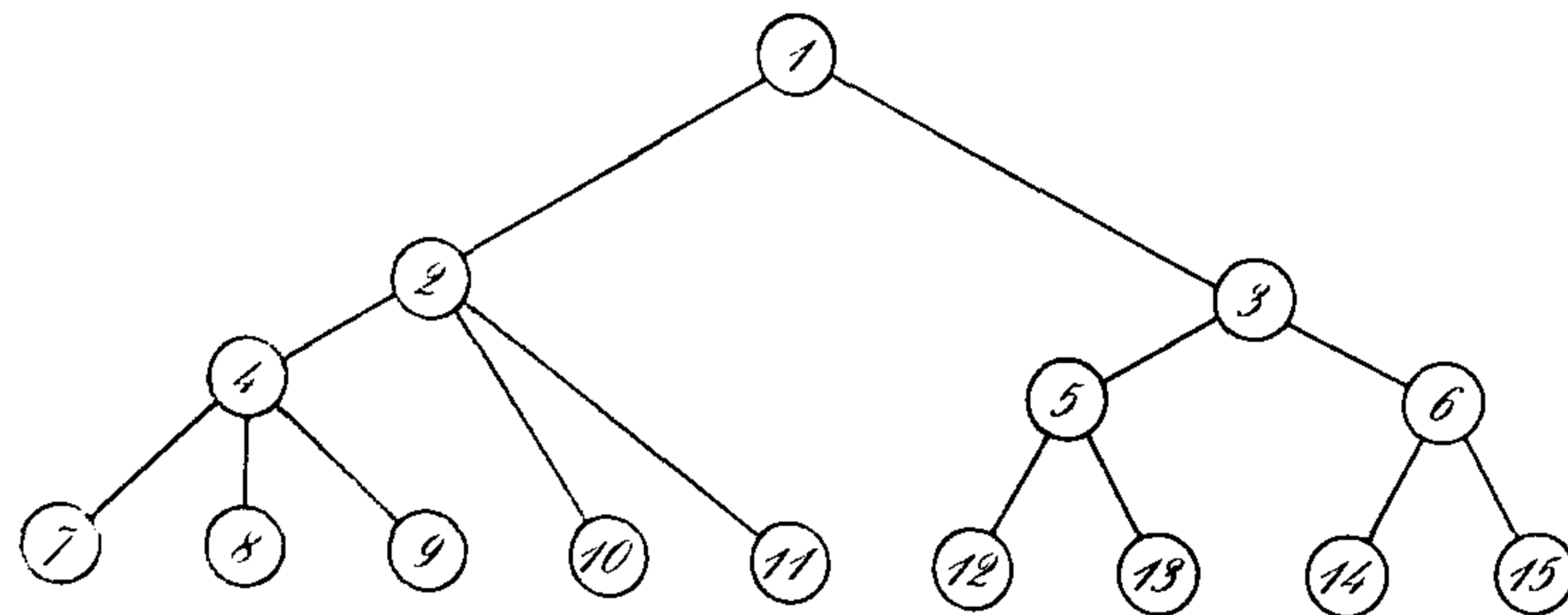


Рис. 7.4

Назовем исполнительный элемент нормально функционирующим, если он работоспособен сам и нормально функционируют все управляющие им элементы, т. е. имеется цепочка связи этого элемента до управляющего элемента самого высокого уровня.

Пусть x_i есть индикаторная функция элемента i :

$$x_i = \begin{cases} 1, & \text{если } i\text{-й исполнительный элемент функционирует нормально,} \\ 0, & \text{если } i\text{-й исполнительный элемент не функционирует нормально,} \end{cases}$$

а δ_i есть «вклад» элемента i . Это означает, что если все исполнительные элементы нормально функционируют, то выходной эффект системы равен $\sum_{i=1}^n \delta_i$. Сами величины δ_i могут иметь любую размерность. Тогда можно записать случайное значение суммарного эффекта всех элементов как $\sum_{i=1}^n x_i \delta_i$. Для каждого исполнительного элемента можно представить вероятность нормального функционирования в виде

$$P_i = \Pr\{x_i = 1\} = E\{x_i\} = p_i R_i;$$

здесь p_i — вероятность работоспособности i -го исполнительного элемента, а R_i — вероятность того, что работоспособны все элементы, управляющие им, — она равна

$$R_i = \prod_{j \in G_i} p_j,$$

где G_i — цепочка управляющих элементов непосредственно для i -го элемента.

Коэффициент эффективности системы с аддитивным выходным эффектом может быть записан как

$$W_{\text{сист}} = E\left\{\sum_{i=1}^n x_i \delta_i\right\} = \sum_{i=1}^n E\{x_i\} \delta_i = \sum_{i=1}^n p_i R_i \delta_i. \quad (7.9)$$

Здесь мы воспользовались тем, что математическое ожидание суммы случайных величин равно сумме их математических ожиданий.

ПРИМЕР 7.3. Для системы, структура которой изображена на рис. 7.4, коэффициент эффективности может быть записан в виде

$$W_{\text{сист}} = p_1 p_2 p_4 (p_7 \delta_7 + p_8 \delta_8 + p_9 \delta_9) + p_1 p_2 (p_{10} \delta_{10} + p_{11} \delta_{11} + p_{12} \delta_{12}) + p_1 p_3 p_6 (p_{13} \delta_{13} + p_{14} \delta_{14} + p_{15} \delta_{15}).$$

7.4.2. Система с резервированием исполнительных элементов

Рассмотрим изображенную на рис. 7.5 систему с симметричной ветвящейся структурой, у которой элементы одного и того же уровня абсолютно идентичны между собой.

Пусть считается, что система выполняет требуемую операцию, если нормально функционирует хотя бы один ее исполнительный элемент. В данном случае обычная формула для нагруженного резервирования не годится, поскольку она была получена в предположении независимости исполнительных элементов.

Для каждой группы из a_n исполнительных элементов можно выразить вероятность успешного выполнения задачи с учетом того, что соответствующий элемент второго уровня нормально функционирует, а $q = 1 - p$, как

$$R_{n-1} = p_{n-1}(1 - q^{a_n}). \quad (7.10)$$

Продолжая аналогичные рассуждения, можно, обозначив $Q = 1 - R$, записать для следующего уровня

$$R_{n-2} = p_{n-2}[1 - (Q_{n-1})^{a_{n-1}}]. \quad (7.11)$$

В результате вероятность успешного выполнения системой операции W :

$$W = p_0[1 - Q_1^{a_1}], \quad (7.12)$$

где величины Q находятся рекуррентно указанным выше способом. Таким образом, вычисляя последовательно (7.10), (7.11) и т. д., мы приходим к финальному значению, определяемому согласно (7.12).

ПРИМЕР 7.4. Рассмотрим систему, представляющую собой стойку оборудования, включающую четыре съемных блока, в каждом из которых содержится по шесть съемных плат, из которых каждые три имеют общий интер-

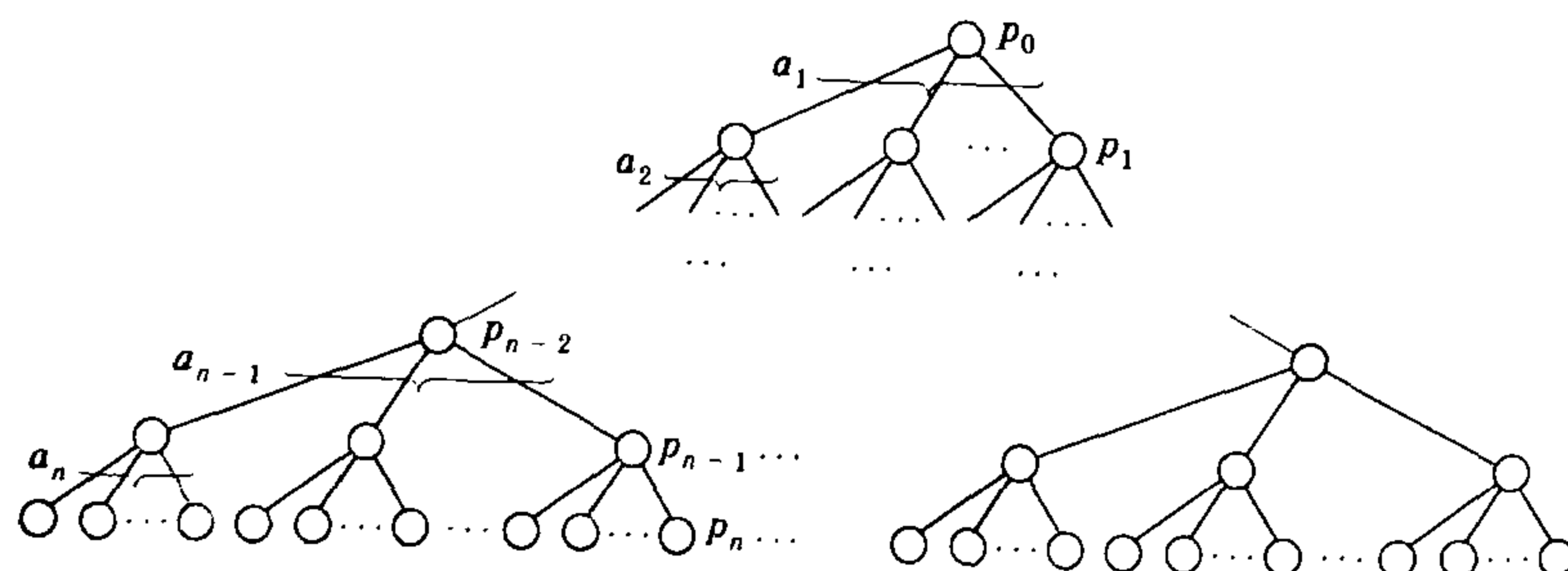


Рис. 7.5

Таблица 7.1

Наименование элемента	Вероятность безотказной работы
Съемная плата	p_1
Интерфейс для трех плат	p_2
Управляющая плата на блок	p_3
Общий блок питания	p_4

фейс. Каждый съемный блок имеет управляющую плату на все шесть съемных плат. Вся стойка в целом имеет один общий блок питания.

Найти вероятность того, что хотя бы одна плата будет функционировать нормально. Соответствующие вероятности безотказной работы каждого из элементов представлены в табл. 7.1.

Решение. На рис. 7.6, а показана конструкция системы, а на рис. 7.6, б — ее формализованная структура.

Формулы для расчета имеют вид

$$R_3 = p_3[1 - (1 - R_4)^3], R_2 = p_2[1 - (1 - R_3)^2], W = R_1 = p_1[1 - (1 - R_2)^4].$$

Если, например, положить все $p_i = 0.9$, то получим

$$R_3 = 0.9 \cdot [1 - (1 - 0.9)^3] \approx 0.899,$$

$$R_2 = 0.9 \cdot [1 - (1 - 0.899)^2] \approx 0.891,$$

$$R_1 = 0.9 \cdot [1 - (1 - 0.891)^4] \approx 0.899,$$

т. е. видно, что фактически надежность такой системы определяется единственным незарезервированным элементом — элементом верхнего уровня иерархии.

Можно заметить, что в данном случае мы имеем обычную приводимую параллельно-последовательную структуру.

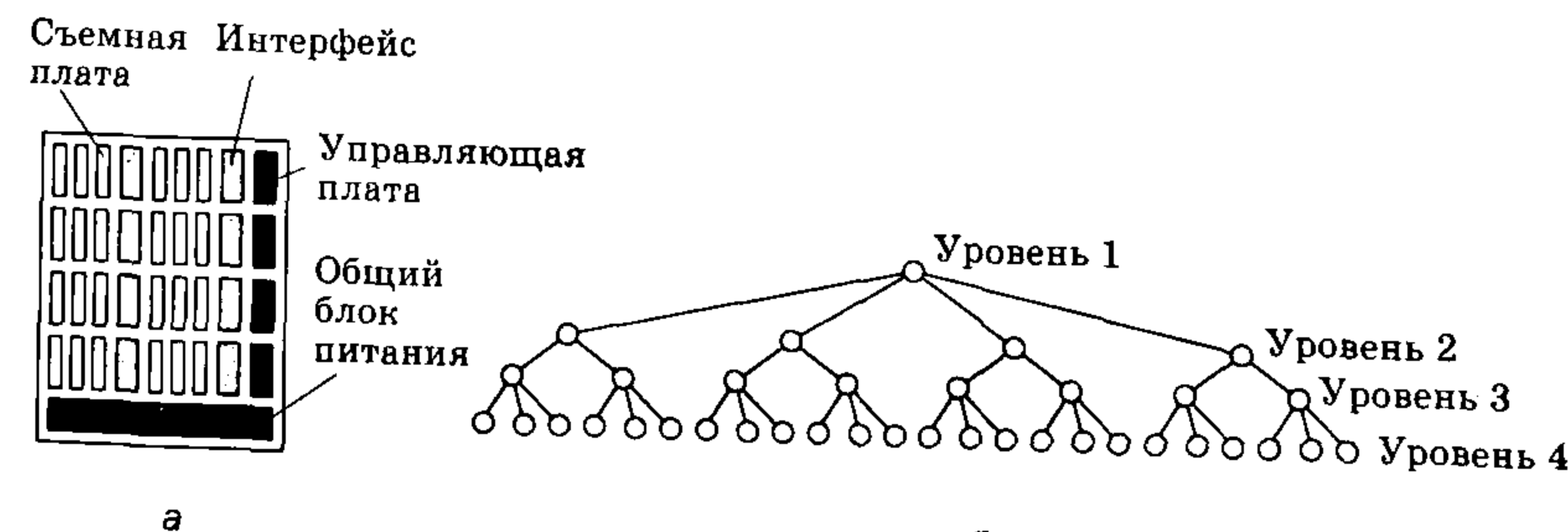


Рис. 7.6

7.4.3. Система с общим видом показателя эффективности

Рассмотрим симметричную ветвящуюся систему, у которой все элементы одного и того же уровня идентичны, а коэффициенты ветвления в вершинах одного и того же уровня также равны (рис. 7.5). Пусть p_k есть вероятность работоспособного состояния элемента, принадлежащего вершине k -го уровня, а a_{k+1} — коэффициент ветвления вершины этого же уровня.

Пусть эффективность функционирования ветвящейся системы $\Phi(x)$ зависит от числа нормально функционирующих элементов нижнего, n -го, уровня иерархии x . Обозначим через $H(x)$ вероятность того, что нормально функционируют x элементов нижнего уровня. Тогда коэффициент эффективности ветвящейся системы определится как

$$W = E\{\Phi(x)\} = \sum_{x=0}^{N_n} \Phi(x) H(x), \quad (7.13)$$

где N_n — полное число исполнительных элементов (нижнего уровня):

$$N_n = \prod_{k=0}^{n-1} a_k. \quad (7.14)$$

Во многих случаях функция $\Phi(x)$ может быть представлена (или аппроксимирована) полиномом конечной и, как правило, невысокой степени m :

$$\Phi(x) = \sum_{k=0}^m c_k x^k. \quad (7.15)$$

В этом случае эффективность системы можно записать в виде

$$W = E\{\Phi(x)\} = E\left\{ \sum_{k=0}^m c_k x^k \right\} = \sum_{k=0}^m c_k M_n^k, \quad (7.16)$$

т. е. задача сводится к нахождению начальных моментов M_n^k распределения $H(x)$. Найдем рекуррентное выражение для этих начальных моментов.

Обозначим через $P_k(z)$ вероятность того, что на k -м уровне нормально функционирует z элементов. Тогда моментная производящая функция может быть записана как

$$\varphi_k(s) = \sum_{z=0}^{N_k} P_k(z) e^{sz}. \quad (7.17)$$

Допустим, что на $(k-1)$ -м уровне нормально функционируют x элементов. Тогда для того, чтобы на k -м уровне нормально функционировало z элементов, необходимо, чтобы в ветвях нормально функционирующих элементов $(k-1)$ -го уровня отказало $(a_k x - z)$ элементов k -го уровня. Окончательно по формуле полной вероятности получаем

$$P_k(z) = \sum_{x=\lceil z/a_k \rceil}^{N_{k-1}} P_{k-1}(x) \binom{a_k x}{z} p_k^z q_k^{a_k x - z}. \quad (7.18)$$

здесь через $\lceil z/a_k \rceil$ обозначена целая часть числа z/a_k .

Теперь для моментной производящей функции можно записать

$$\varphi_k(s) = \sum_{z=0}^{N_k} \sum_{x=\lceil z/a_k \rceil}^{N_{k-1}} P_{k-1}(x) \binom{a_k x}{z} p_k^z q_k^{a_k x - z} e^{sz}. \quad (7.19)$$

Изменив порядок суммирования и проведя ряд несложных преобразований, получаем

$$\begin{aligned} \varphi_k(s) &= \sum_{x=0}^{N_{k-1}} \sum_{z=0}^{a_k x} P_{k-1}(x) \binom{a_k x}{z} q_k^{a_k x - z} (p_k e^s)^z = \\ &= \sum_{x=0}^{N_{k-1}} P_{k-1}(x) \sum_{z=0}^{a_k x} \binom{a_k x}{z} q_k^{a_k x - z} (p_k e^s)^z = \\ &= \sum_{x=0}^{N_{k-1}} P_{k-1}(x) (q_k + p_k e^s)^{a_k x}. \end{aligned} \quad (7.20)$$

Окончательно получаем рекуррентное выражение вида

$$\varphi_k(s) = \varphi_{k-1}((p_k e^s + q_k)^{a_k}). \quad (7.21)$$

Это рекуррентное выражение позволяет вычислять начальные моменты распределения числа нормально функционирующих исполнительных элементов. В частности, для первых двух моментов можно записать:

$$M_n^{(1)} = M_{n-1}^{(1)} a_n p_n = \dots = p_0 \prod_{k=1}^n a_k p_k.$$

$$\begin{aligned} M_n^{(2)} &= M_{n-1}^{(2)} (a_n p_n)^2 + M_{n-1}^{(1)} a_n p_n q_n = \dots = \\ &= p_0 \prod_{k=1}^n a_k p_k \left(\prod_{k=1}^n a_k p_k + \prod_{k=1}^n q_k \prod_{i=k+1}^n a_i p_i \right). \end{aligned}$$

Моменты более высоких порядков имеют громоздкие выражения, однако их вычисление с помощью компьютера не представляет труда.

§ 7.5. СИСТЕМА С ПЕРЕКРЫВАЮЩИМИСЯ ЗОНАМИ ДЕЙСТВИЯ

Часто встречаются территориально-распределенные системы, у которых соседние зоны действия частично перекрываются. Это могут быть сотовые системы связи, системы ПВО и т. п.

Рассмотрим для примера систему, состоящую из четырех подсистем Z_i , каждая из которых обеспечивает работу в своей зоне (рис. 7.7). Район действия системы в целом определяется как

$$\Omega = \cup Z_i.$$

На рисунке в каждой зоне выделены области, перекрываемые двумя или даже тремя соседними зонами: темно-серым закрашены области пересечения двух зон, а черным — район пересечения трех зон. Та часть зоны Z_i , которая не перекрывается ни с одной другой, закрашена светло-серым. Часть зоны, не пересекающаяся ни с одной другой, будет обозначаться через S_i , район пересечения двух зон — через S_{ij} , а район пересечения трех зон — через S_{ijk} . Таким образом, все упомянутые районы S не пересекаются, и район действия системы в целом есть объединение всех этих непересекающихся районов. Площадь всех районов S будем обозначать через s .

Соответствующие коэффициенты эффективности выделенных областей обозначим, как и ранее, через W с соответствующими индексами. Если, например, откажет элемент, ответственный за зону Z_2 , коэффициенты эффективности областей, покрываемых этой зоной, также изменятся: вместо W_{12} и W_{24} получим соответственно W_1 , W_4 и т. п. Все коэффициенты эффективности для областей пересечения определяются индивидуально в каждой практической задаче.

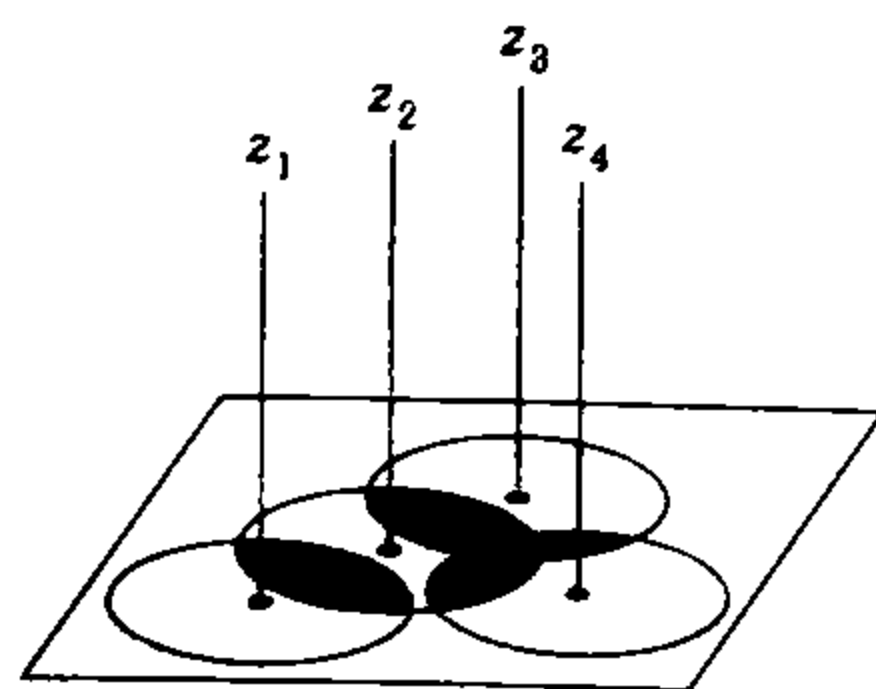


Рис. 7.7

Такое разбиение области действия системы из n элементов на подобласти приводит в результате к числу областей, меньшему чем 2^n , т. е. упрощает вычисление коэффициента эффективности системы.

В результате подсчет коэффициента эффективности системы сводится к вычислению коэффициентов эффективности каждой из областей с уче-

том их площади и их взвешенному суммированию. Составляющая в результирующей сумме для области S_i без перекрытия S с другими зонами равняется

$$d_i = s_i W_i p_i. \quad (7.22)$$

Составляющая для области, образованной перекрытием зон Z_i и Z_j , равна

$$d_{ij} = s_{ij} [W_{ij} p_i p_j + W_i p_i q_j + W_j q_i p_j]. \quad (7.23)$$

Составляющая для области, образованной перекрытием зон Z_i , Z_j и Z_k , равна

$$d_{ijk} = s_{ijk} [W_{ijk} p_i p_j p_k + W_{jk} q_i p_j p_k + W_{ik} p_i q_j p_k + W_{ij} p_i p_j q_k + W_{kj} q_i q_j p_k + W_{ji} q_i p_j q_k + W_{ik} p_i q_j q_k] \quad (7.24)$$

и т. д.

В высоконадежных системах появление более одного отказа маловероятно, поэтому можно написать приближенную формулу потерь в зоне Z_k в случае отказа ее основного элемента: полные потери будут в области S_k , а в остальных областях перекрытия с другими зонами будет соответствующее снижение эффективности функционирования.

Иными словами, с вероятностью q_k потери в k -й зоне характеризуются величиной Q , равной

$$Q_k = s_k W_k + s_{kj} (W_{kj} - W_k) + s_{kji} (W_{kji} - W_k) + \dots \quad (7.25)$$

Результирующая приближенная формула записывается в виде

$$Q_{\text{сист}} \approx \sum_{k=1}^n q_k Q_k. \quad (7.26)$$

Безусловно, этими примерами рассмотрение систем с перекрывающимися зонами действия не ограничивается.

Глава 8

Оптимальное резервирование

Рассматривается проблема распределения ресурсов для оптимального обеспечения систем резервными элементами. Предлагаются строгие и приближенные методы решения этой задачи.

§ 8.1. ФОРМУЛИРОВКА ПРОБЛЕМЫ

Одним из эффективных методов повышения надежности сложных систем является резервирование. Однако всегда возникают вопросы. Что резервировать в первую очередь? Какой резерв необходим? Как рациональнее всего распределить ресурсы повышения надежности между подсистемами? Последний вопрос представляется наиболее актуальным.

Возможны две основных постановки задачи оптимального резервирования.

- (1) Путем резервирования добиться максимально возможного значения выбранного показателя надежности системы при заданных ограничениях на общие затраты, связанные с введением резервных элементов.
- (2) Путем резервирования достичь требуемого значения показателя надежности системы при минимально возможных затратах на резервные элементы.

8.1.1. Математическая формулировка задачи оптимального резервирования при одном ограничивающем факторе

Пусть имеется система, состоящая из n независимых элементов. Надежность каждого i -го элемента может быть повышена за счет использования резервных элементов i -го типа (рис. 8.1). Обозначим полное число резервных элементов в i -й резервной группе через x_i .

Рассмотрим показатели надежности типа времени безотказной работы (ВБР) или коэффициента готовности (КГ), которые можно представить в виде произведения соответствующих вероятностей для отдельных подсистем. Обозначим этот обобщенный показатель надежности через $R_i(x_i)$, где x_i —

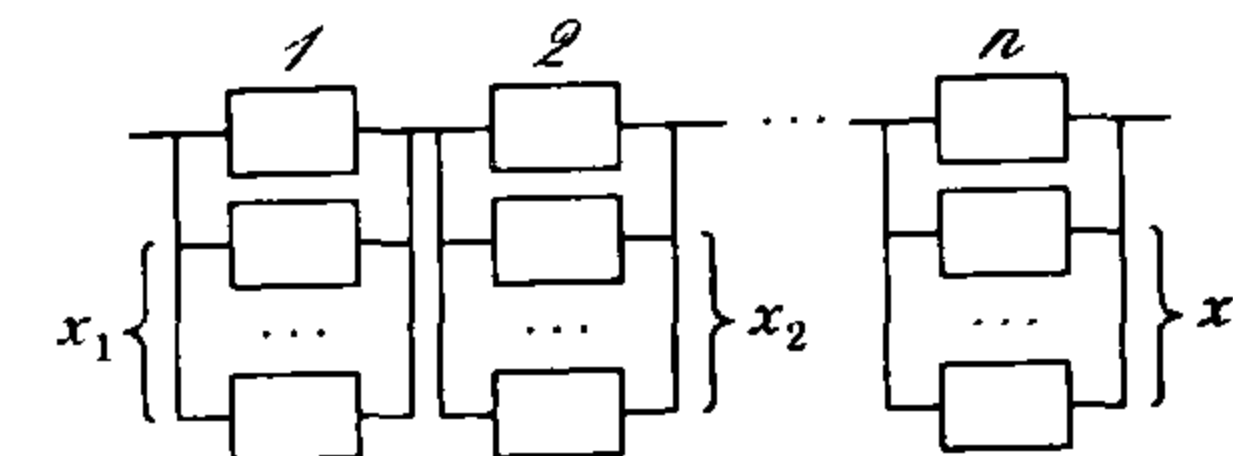


Рис. 8.1

общее число элементов i -й резервной группы. Функция $R_i(x_i)$ зависит от того, что представляет из себя подсистема, какова степень нагруженности резерва и от других факторов. Различные функции такого рода рассмотрены в предыдущих главах.

Пусть известна также функция стоимости этих элементов $C_i(x_i)$. Обычно полагают, что эта функция во всех прикладных задачах линейно зависит от числа элементов:

$$C_i(x_i) = \sum_{i=1}^n c_i x_i; \quad (8.1)$$

здесь c_i — стоимость одного элемента i -го типа*.

Иногда в качестве ограничивающего фактора рассматривается не стоимость, а вес или объем резервных элементов, особенно если проектировщик имеет дело с такими объектами, как космические корабли или подводные лодки. Возможно рассмотрение и нескольких ограничивающих факторов совместно.

Задача (1) может быть записана в виде

$$\max_{1 \leq i \leq n} \left\{ \prod_{i=1}^n R_i(x_i) \mid \sum_{i=1}^n C_i(x_i) \leq C_{\text{доп}} \right\}, \quad (8.2)$$

где $C_{\text{доп}}$ есть допустимое (заданное) ограничение на суммарную стоимость резервных элементов.

Задача (2) — обратная — формулируется в виде

$$\min_{1 \leq i \leq n} \left\{ \sum_{i=1}^n C_i(x_i) \mid \prod_{i=1}^n R_i(x_i) \geq R_{\text{треб}} \right\}, \quad (8.3)$$

где $R_{\text{треб}}$ есть требуемое значение показателя надежности системы.

* Линейная функция стоимости соответствует одноразовому созданию комплекта резервных элементов. Она действительно может использоваться вместе с показателем типа ВБР. Однако при наличии восстановления отказавших элементов или их восполнения из внешнего источника, когда в качестве показателя надежности выбирается стационарный коэффициент готовности, использование такой функции стоимости вряд ли экономически корректно.

8.1.2. Математическая формулировка задачи оптимального резервирования при нескольких ограничивающих факторах

Обычно на практике рассматривается задача типа

$$\max \left\{ \prod_{i=1}^n R_i(x_i) \mid \sum_{i=1}^n C_i^{(1)}(x_i) \leq C_{\text{доп}}^{(1)}, \dots, \sum_{i=1}^n C_i^{(M)}(x_i) \leq C_{\text{доп}}^{(M)} \right\}, \quad (8.4)$$

где верхний индекс показывает тип (номер) ограничивающего фактора. Как уже упоминалось, такими различными факторами могут быть стоимость, вес, габариты и пр.

Обратная задача при нескольких ограничениях не имеет столь простой формулировки, как в случае единственного ограничения. Обычно рассматривается решение, при котором достигается требуемое значение показателя надежности системы, а остальные факторы находятся в «приемлемой зоне». Например, если ограничивающими факторами являются стоимость и вес резервных элементов, а сами эти два фактора противоречивы, т. е. уменьшение веса ведет к увеличению стоимости, то одно и то же значение показателя надежности системы может быть достигнуто либо при очень большом суммарном весе резерва, либо при слишком большой его стоимости. Например, можно взять больше тяжелых, но дешевых элементов, проиграв в весе системы, а можно взять легкие, но дорогие элементы, проиграв в стоимости. Окончательный выбор зависит от характера системы и ее предназначения. Так, для наземной системы может быть более критической стоимость, а для подлодки — суммарный объем или вес резервных элементов.

В этих случаях возникает необходимость поиска компромиссного решения, т. е. задача сводится к многокритериальной оптимизации: нахождению множества Парето. Формально эту задачу можно записать в виде

$$\min \left\{ \sum_{i=1}^n C_i^{(1)}(x_i), \dots, \sum_{i=1}^n C_i^{(M)}(x_i) \mid \prod_{i=1}^n R_i(x_i) \geq R_{\text{треб}} \right\}. \quad (8.5)$$

Будем обозначать через MIN прописными буквами оптимизацию по Парето. Такая запись нуждается в пояснении. Требуемый уровень показателя надежности может быть достигнут несколькими путями. Например, несколько различных векторов $X^{(1)} = (x_1^{(1)}, \dots, x_n^{(1)})$, $X^{(2)} = (x_1^{(2)}, \dots, x_n^{(2)})$, ..., $X^{(K)} = (x_1^{(K)}, \dots, x_n^{(K)})$ обеспечивают выполнение условия $\prod_{i=1}^n R_i(x_i) \geq R_{\text{треб}}$, давая при этом различные набо-

ры $\sum_{i=1}^n C_i^{(1)}(x_i), \dots, \sum_{i=1}^n C_i^{(M)}(x_i)$ значений ограничивающих факторов. Все эти решения и образуют множество Парето. В двумерном случае, например, для стоимости C и веса W множество Парето образует некоторую оболочку — рис. 8.2*.

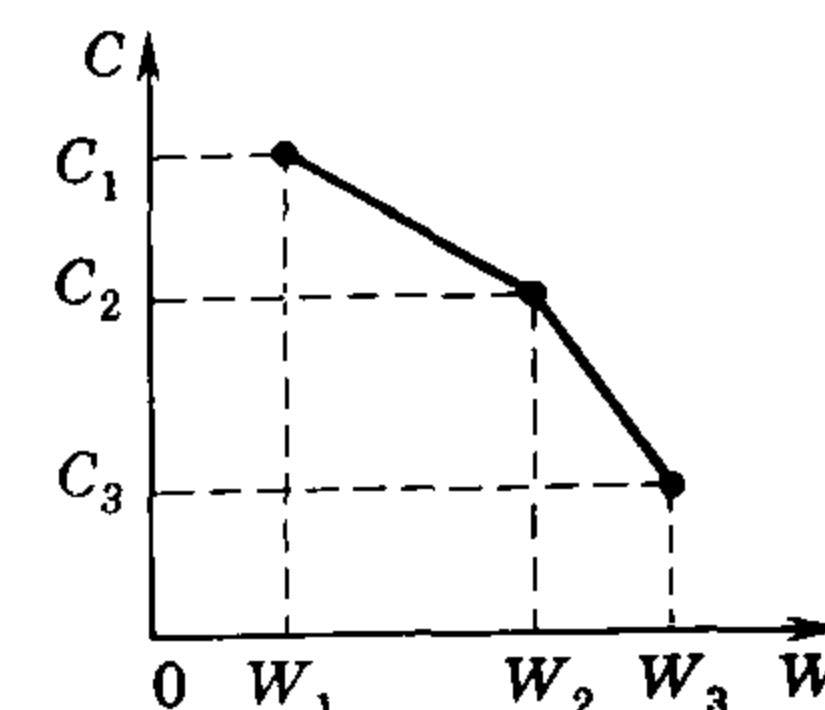


Рис. 8.2

Каждое из решений характеризуется своим уникальным набором значений C и W . Фактически на этом математическая часть задачи заканчивается. Наступает момент, когда вступает в силу принимающее решение лицо, которое выбирает лучший вариант. Это дело неформальное и очень персональное — здесь вступают в силу иногда самые неожиданные качественные соображения, базирующиеся на опыте, интуиции, а иногда даже и на вкусе этого лица.

§ 8.2. МЕТОД МНОЖИТЕЛЕЙ ЛАГРАНЖА

Первые попытки решения задачи оптимального резервирования опирались на метод неопределенных множителей Лагранжа. Вообще говоря, использование этого метода в данном случае некорректно, поскольку, во-первых, все переменные в задаче оптимального резервирования принимают дискретные значения, а во-вторых, ограничение имеет вид неравенств, а не строгих равенств. Тем не менее этот метод интересен и важен, поскольку он в ряде случаев позволяет быстро находить оценочные значения.

Продemonстрируем этот метод на примере высоконадежной последовательной системы, повышение надежности элементов которой возможно с использованием *горячего* резервирования (рис. 8.1). Целевую функцию $R(X)$ для такой системы можно записать приближенно как

$$R(X) = \prod_{i=1}^n R_i(x_i) \approx 1 - \sum_{i=1}^n Q_i(x_i), \quad (8.6)$$

* Множество Парето-оптимальных решений характеризуется тем, что для любого из них не существует ни одного решения, которое было бы заведомо лучше данного, т. е. давало бы лучшее значение хотя бы одного из выбранных показателей и не худшие значения всех остальных показателей.

Парето В. (Pareto V., 1848—1923) — итальянский социолог и экономист.

где $X = (x_1, \dots, x_n)$ — набор резервных элементов всех типов, $R_i(x_i)$ — показатель надежности i -й резервной группы, а $Q_i(x_i) = 1 - R_i(x_i)$. Показателем надежности $R_i(x_i)$ может в данном случае служить вероятность безотказной работы или коэффициент готовности, т. е. показатель, который выражается в виде произведения соответствующих показателей для последовательной системы.

Заменим задачу максимизации функции $R(X)$ задачей минимизации функции $Q(X) = 1 - R(X)$. В случае горячего резерва с x_i резервными элементами $Q_i(x_i) = q_i^{x_i+1}$.

Поставим обратную задачу оптимального резервирования: найти вектор состава резервных элементов $X = (x_1, \dots, x_n)$, доставляющий

$$\min_X \left\{ \sum_{i=1}^n c_i x_i \mid \sum_{i=1}^n q_i^{x_i+1} = Q^0 \right\}, \quad (8.7)$$

где Q^0 есть требуемое значение.

Для решения задачи методом неопределенных множителей Лагранжа нужно записать функцию Лагранжа $\tilde{\lambda}(X)$ в виде

$$\tilde{\lambda}(X) = \sum_{i=1}^n c_i x_i + \tilde{\lambda} \sum_{i=1}^n q_i^{x_i+1}, \quad (8.8)$$

где $\tilde{\lambda}$ — неопределенный множитель Лагранжа*, а затем записать систему $(n+1)$ -го уравнений:

$$\frac{\partial}{\partial x_i} \left(\sum_{i=1}^n c_i x_i + \tilde{\lambda} \sum_{i=1}^n q_i^{x_i+1} \right) = 0, \quad i = 1, 2, \dots, n, \quad (8.9)$$

$$\sum_{i=1}^n q_i^{x_i+1} = Q^0. \quad (8.10)$$

Уравнение (8.10) называется уравнением связи.

В результате дифференцирования получаем для первых n уравнений

$$c_i + \tilde{\lambda} (\ln q_i) q_i^{x_i+1} = 0, \quad (8.11)$$

откуда следует

$$q_i^{x_i+1} = -\frac{c_i}{\tilde{\lambda}_i \ln q_i}. \quad (8.12)$$

* Мы использовали стандартное обозначение неопределенного множителя Лагранжа, снабдив его знаком *тильда*. Не путать с интенсивностью отказов!

Обозначив $\alpha_i = -c_i / \ln q_i$, после подстановки (8.12) в уравнение связи (8.10) находим

$$\tilde{\lambda} = -\frac{1}{Q^0} \sum_{i=1}^n \alpha_i. \quad (8.13)$$

Решение получается путем подстановки (8.13) в (8.12):

$$q_i^{x_i+1} = Q^0 \frac{\alpha_i}{\sum_{i=1}^n \alpha_i}, \quad (8.14)$$

откуда окончательно имеем

$$x_i = \frac{1}{\ln q_i} \ln \left(Q^0 \frac{\alpha_i}{\sum_{i=1}^n \alpha_i} \right). \quad (8.15)$$

Заметим, что в других предположениях о характере резерва получить какие-либо формулы в замкнутом виде не удастся.

Метод множителей Лагранжа рассмотрен нами, поскольку он помогает понять важный принцип, на котором базируется метод наискорейшего спуска, широко используемый в инженерной практике.

Физически пояснить суть метода неопределенных множителей достаточно просто, рассмотрев процесс численного нахождения решения задачи оптимального резервирования. Из (8.9) следует, что

$$\frac{\partial}{\partial x_i} \sum_{i=1}^n q_i^{x_i+1} / \frac{\partial}{\partial x_i} \sum_{i=1}^n c_i x_i = \tilde{\lambda} = \text{const} \quad (8.16)$$

для всех $i = 1, 2, \dots, n$. Решение состоит в одновременном увеличении всех x_i таким образом, чтобы при этом выполнялось условие (8.16). Этот процесс продолжается до тех пор, пока не будет выполнено условие (8.10).

§ 8.3. МЕТОД ПОКООРИНАТНОГО НАИСКОРЕЙШЕГО СПУСКА

8.3.1. Общее описание метода

Во многих прикладных задачах с успехом используется метод оптимизации, называемый методом наискорейшего спуска (подъема). Этот метод может быть успешно применен для решения проблемы оптимального резервирования, поскольку функция $R(X)$

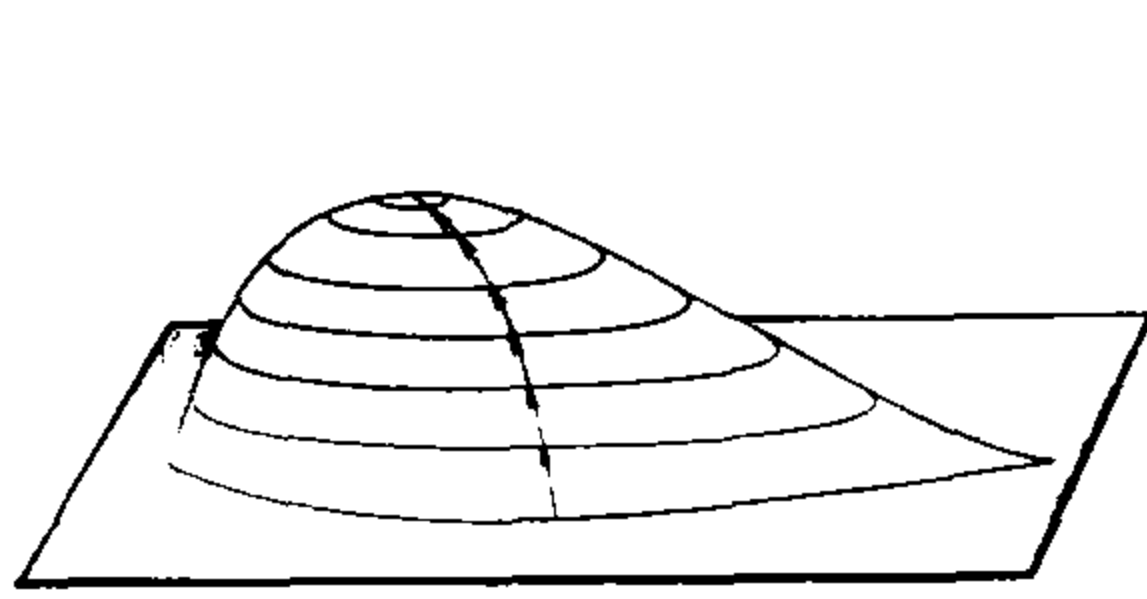


Рис. 8.3

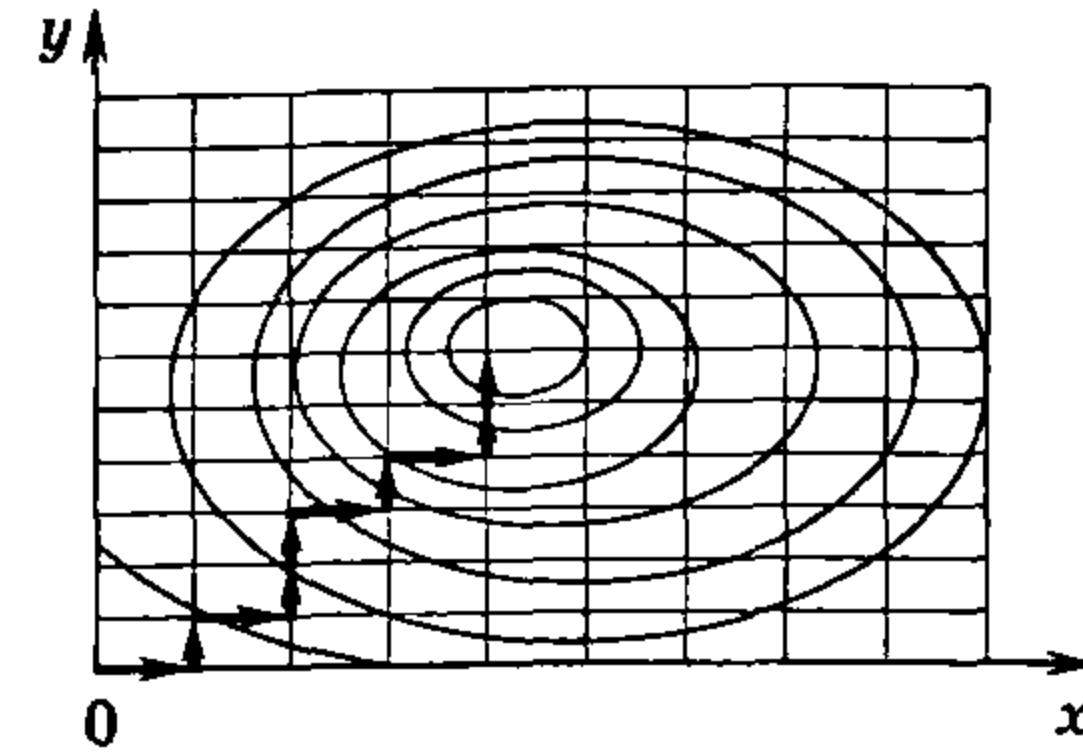


Рис. 8.4

унимодальна, т. е. не имеет локальных минимумов (максимумов). Пример такой функции представлен на рис. 8.3.

Идею метода лучше всего пояснить примером с наикратчайшим подъемом в тумане на вершину холма из некоторой определенной точки у его основания: каждый шаг вверх выбирается в направлении наиболее крутого подъема, т. е. максимального градиента.

Покоординатный спуск отличается от градиентного спуска тем, что на каждом шаге процесса разрешено движение не в любом направлении, а лишь в направлении одной из осей координат, по которой в данной точке наблюдается наибольшее относительное приращение (рис. 8.4), причем величина шага предполагается не бесконечно малой, а конечной. При этом не обязательно, чтобы длина шага была одинакова по всем направлениям.

8.3.2. Применение метода к задаче оптимального резервирования

Рассмотрим последовательную систему. Пусть исходная система состоит из n групп однотипных элементов, надежность каждой из которых может быть повышена путем введения резервных элементов. Если рассматривается обеспечение системы запасными элементами, то можно считать, что каждая такая группа — это полная совокупность однотипных элементов системы: все элементы i -го типа, имеющиеся в системе, образуют i -ю группу. Если же рассматривается горячий резерв, то группы однотипных элементов могут быть расположены в аппаратуре так, что для каждой из них возможен только индивидуальный резерв; такие группы рассматриваются независимо.

Процедура оптимизации представляет собой многошаговый процесс, где на каждом очередном шаге N проверяется, добавле-

ние какого из резервных элементов дает наибольший относительный прирост в надежности системы.

Обозначим через $R_i(x_i^{(N)})$ показатель надежности i -го участка с $x_i^{(N)}$ резервными элементами после N -го шага процесса наискорейшего спуска, через $C_i(x_i^{(N)})$ стоимость этих резервных элементов, через $R(X^{(N)})$ показатель надежности системы в целом, а через $C(X^{(N)})$ ее стоимость; здесь $X^{(N)} = (x_1^{(N)}, \dots, x_n^{(N)})$. Тогда в предположении независимости отказов элементов на различных участках системы можно записать:

$$R^{(N)} = R(X^{(N)}) = R(x_1^{(N)}, \dots, x_n^{(N)}) = \prod_{i=1}^n R_i(x_i^{(N)}). \quad (8.17)$$

Суммарная стоимость резервных элементов равна

$$C^{(N)} = C(X^{(N)}) = C(x_1^{(N)}, \dots, x_n^{(N)}) = \sum_{i=1}^n c_i x_i^{(N)}. \quad (8.18)$$

Для удобства обозначим через $X_i^{(N)}$ вектор $X^{(N)}$ с исключенной компонентой $x_i^{(N)}$, т. е.

$$X_i^{(N)} = (x_1^{(N)}, \dots, x_{i-1}^{(N)}, \emptyset, (x_{i+1}^{(N)}, \dots, x_n^{(N)}).$$

Будем также иметь в виду, что $X^{(N)} = (X_i^{(N)}, x_i^{(N)})$. На N -м шаге процесса, в соответствии со сформулированным правилом, добавляется резервный элемент с индексом k таким, у которого максимально значение относительного приращения:

$$\begin{aligned} \gamma_k^{(N)}(x_i^{(N)}) &= \max_{1 \leq i \leq n} \frac{R(X_i^{(N)}, x_i^{(N)} + 1) - R(X^{(N)})}{C(X_i^{(N)}, x_i^{(N)} + 1) - C(X^{(N)})} = \\ &= \max_{1 \leq i \leq n} \frac{R(X_i^{(N)}, x_i^{(N)} + 1) - R(X^{(N)})}{c_i}. \end{aligned} \quad (8.19)$$

Заметим, что числитель в (8.19) может быть преобразован следующим образом:

$$R(X_i^{(N)}, x_i^{(N)} + 1) - R(X^{(N)}) = R(X^{(N)}) \left[\frac{R_i(x_i^{(N)} + 1)}{R_i(x_i^{(N)})} - 1 \right]. \quad (8.20)$$

Следовательно, вместо (8.19) индекс k можно искать по условию

$$\gamma_k^{(N)}(x_i^{(N)}) = \max_{1 \leq i \leq n} \frac{R_i(x_i^{(N)} + 1) - R_i(x_i^{(N)})}{c_i R_i(x_i^{(N)})}. \quad (8.21)$$

8.3.3. Алгоритм решения задачи

Процедура решения задачи оптимального резервирования состоит в следующих шагах.

- (1) Перед первым шагом в системе резервных элементов нет, т. е. $x_1^{(0)} = \dots = x_n^{(0)} = 0$ или $X^{(0)} = 0$.
- (2) Вычисляются значения $\gamma_i^{(0)}(x_i^{(0)})$ для всех $i = 1, 2, \dots, n$.
- (3) Находится максимальное из всех значений $\gamma_i^{(0)}(x_i^{(0)})$, пусть оно имеет индекс k .
- (4) Вычисляется новое значение $x_k^{(1)} = x_k^{(0)} + 1$.
- (5) Все остальные $x_i^{(0)}$, $i \neq k$, оставляются без изменений, но меняют верхний индекс: $x_i^{(0)} \rightarrow x_i^{(1)}$.
- (6) Строится новый вектор состава резервных элементов системы $X^{(1)} = (X_k^{(0)}, x_k^{(1)})$.
- (7) Вычисляется значение $R(X^{(1)})$.
- (8) Вычисляется значение $C(X^{(1)})$.
- (9) Вычисляется новое значение $\gamma_k^{(1)}(x_k^{(1)})$.
- (10) Все остальные $\gamma_i^{(0)}(x_i^{(0)})$, $i \neq k$, оставляются без изменений, но меняют верхний индекс: $\gamma_i^{(0)}(x_i^{(0)}) \rightarrow \gamma_i^{(1)}(x_i^{(1)})$.
- (11) GOTO (3).

ПРАВИЛО ОСТАНОВКИ

- Решение прямой задачи оптимального резервирования $X^{(N)}$ достигается на шаге N таком, при котором

$$C(X^{(N)}) \leq C^0 < C(X^{(N+1)}). \quad (8.22)$$

Полученное значение $R(X^{(N)})$ является максимально возможным показателем надежности системы при заданных ограничениях на суммарную стоимость резервных элементов.

- Решение обратной задачи оптимального резервирования $X^{(N)}$ достигается на шаге N таком, при котором

$$R(X^{(N-1)}) < R^0 \leq R(X^{(N)}). \quad (8.23)$$

Полученное значение $C(X^{(N)})$ является минимально возможной суммой затрат на резервные элементы при условии достижения требуемого показателя надежности системы в целом.

ПРИМЕР 8.1. Последовательная система состоит из шести элементов с характеристиками, приведенными в табл. 8.1.

Требуется найти оптимальное число резервных элементов для системы, если время выполнения операции равно $t_0 = 1000$ ч, для двух случаев:

- (1) требование по вероятности безотказной работы системы равно 0.9995;
- (2) допустимые затраты на резервные элементы составляют 70 стоимостных единиц (ст. ед.).

Резерв предполагается *холодным*, что соответствует случаю обеспечения системы запасными элементами.

Решение Находим для всех резервных групп значения соответствующих параметров распределения Пуассона:

$$a_i = \lambda_i n_i t_0:$$

$$\begin{aligned} a_1 &= 1 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.05; & a_2 &= 1 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.1; \\ a_3 &= 1 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.05; & a_4 &= 1 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.1; \\ a_5 &= 8 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.4; & a_6 &= 8 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.8. \end{aligned}$$

Формула расчета вероятности появления x_i отказов в каждой резервной группе в соответствии с распределением Пуассона имеет вид

$$q_i(k) = \frac{(\lambda_i n_i t_0)^k}{k!} \exp(-\lambda_i n_i t_0). \quad (8.24)$$

Значения величин $q_i(x_i)$, полученных с использованием *Microsoft Excel*, сведены в табл. 8.2, из которой следует, что в данном случае действительно выполняется условие $Q_i(x_i) \gg Q_i(x_i + 1)$ при достаточно больших x_i , поэтому возможно использовать приближение

$$\gamma_i(x_i) \approx \frac{q_i(x_i)}{c_i}. \quad (8.25)$$

Таблица 8.1

Номер элемента	Число элементов	Интенсивность отказов, λ_i (10^{-5} 1/ч)	Стоимость элемента, c_i
1	5	1	1
2	10	1	1
3	5	1	8
4	10	1	8
5	5	8	1
6	10	8	1

Таблица 8.2

x_i	$q_1(x_1)$	$q_2(x_2)$	$q_3(x_3)$	$q_4(x_4)$	$q_5(x_5)$	$q_6(x_6)$
0	0.0476	0.0905	0.0476	0.0905	0.268	0.359
1	0.00119	0.00452	0.00119	0.00452	0.0536	0.144
2	1.98E-05	0.000151	1.98E-05	0.000151	0.00715	0.0383
3	2.48E-07	3.77E-06	2.48E-07	3.77E-06	0.000715	0.00767
4	2.48E-09	7.54E-08	2.48E-09	7.54E-08	5.72E-05	0.00123
5	2.06E-11	1.26E-09	2.06E-11	1.26E-09	3.81E-06	0.000164
6	1.47E-13	1.8E-11	1.47E-13	1.8E-11	2.18E-07	1.87E-05
7	9.22E-16	2.24E-13	9.22E-16	2.24E-13	1.09E-08	1.87E-06
8	5.12E-18	2.49E-15	5.12E-18	2.49E-15	4.84E-10	1.66E-07
9	2.56E-20	2.49E-17	2.56E-20	2.49E-17	1.94E-11	1.33E-08

* Здесь и далее запись E – MN означает 10^{-MN} , например $1.98E-05 = 1.98 \cdot 10^{-5}$; $1.94E-11 = 1.94 \cdot 10^{-11}$; соответственно $E8 = 10^8$.

Используя этот факт, на базе табл. 8.2 составим табл. 8.3 относительных приращений. В ней в правых верхних углах ячеек проставлены номера соответствующих шагов покоординатного наискорейшего спуска. Остановим решение на 22-м шаге: соответствующие ячейки закрашены серым цветом. Решением задачи на этом шаге является следующий набор резервных элементов: $x_1 = 3$; $x_2 = 4$; $x_3 = 3$; $x_4 = 3$; $x_5 = 5$; $x_6 = 6$.

Решение обратной задачи. Используя табл. 8.2, находим, что ненадежность системы, равная сумме значений в соответствующих ячейках этой таблицы, есть

$$Q^{(22)} = 1.98 \cdot 10^{-5} + 3.77 \cdot 10^{-6} + 1.98 \cdot 10^{-5} + \\ + 5.72 \cdot 10^{-6} + 1.98 \cdot 10^{-5} + 1.64 \cdot 10^{-4} = 2.68 \cdot 10^{-4},$$

что удовлетворяет требованиям, сформулированным в условии задачи. Можно убедиться, что, используя описанную процедуру, улучшить решение нельзя.

Решение прямой задачи. Теперь нам нужно найти оптимальное решение для случая (2) — составим табл. 8.4 возрастания стоимости резервных групп по мере увеличения числа запасных элементов. Выберем опять в табл. 8.3 остановку на 22-м шаге. Проверим, чему равна сумма стоимостей, стоящих в соответствующих ячейках табл. 8.4. Подсчитав, находим значение, равное 76 ст. ед., что превышает допустимое значение. Проверяем, какой элемент был добавлен на последнем 22-м шаге: им оказывается 3-й элемент. Исключаем его и получаем в результате 68 ст. ед., т. е. решение $x_1 = 3$; $x_2 = 4$; $x_3 = 2$; $x_4 = 5$; $x_5 = 7$.

Таблица 8.3

x_i	$\gamma_1(x_1)$	$\gamma_2(x_2)$	$\gamma_3(x_3)$	$\gamma_4(x_4)$	$\gamma_5(x_5)$	$\gamma_6(x_6)$
1	№2 0.0464	№1 0.086	№8 0.0058	№5 0.0107	№4 0.0268	№3 0.027
2	№11 0.00117	№9 0.00437	№15 0.000146	№14 0.000547	№7 0.00581	№6 0.0132
3	№17 1.96E-05	№14 0.000147	№22 2.45E-06	№18 1.84E-05	№13 0.000804	№10 0.003834
4	№26 2.45E-07	№21 3.69E-06	№24 3.07E-08	№24 4.62E-07	№16 8.22E-05	№12 0.000805
5	2.46E-09	7.41E-08	3.07E-10	9.27E-09	№20 6.67E-06	№15 0.000133
6	2.05E-11	1.24E-09	2.56E-12	1.55E-10	№25 4.49E-07	№19 1.81E-05
7	1.47E-13	1.77E-11	1.83E-14	2.22E-12	2.59E-08	№23 2.10E-06
8	9.16E-16	2.22E-13	1.15E-16	2.77E-14	1.30E-09	№27 2.13E-07
9	5.09E-18	2.47E-15	6.37E-19	3.09E-16	5.81E-11	1.91E-08

Таблица 8.4

x_i	$C_1(x_1)$	$C_2(x_2)$	$C_3(x_3)$	$C_4(x_4)$	$C_5(x_5)$	$C_6(x_6)$
0	0	0	0	0	0	0
1	1	1	8	8	1	1
2	2	2	16	16	2	2
3	3	3	24	24	3	3
4	4	4	32	32	4	4
5	5	5	40	40	5	5
6	6	6	48	48	6	6
7	7	7	56	56	7	7
8	8	8	64	64	8	8
9	9	9	72	72	9	9

8.3.4. Приближенное решение

Если построить таблицы с пошаговым решением (см. табл. 8.3), то можно заметить очевидный факт: на любом шаге процесса по координатного наискорейшего спуска для всякого текущего шага N существует величина γ^* такая, что при всех $i = 1, 2, \dots, n$ выполняется условие

$$\gamma_i^{(N)} \leq \gamma^* \leq \gamma_i^{(N+1)}. \quad (8.26)$$

Выполнение этого условия иллюстрируется табл. 8.3, где значения $\gamma_i^{(N)}$, расположенные в ячейках, закрашенных серым цветом, меньше любых значений, расположенных выше, и больше значений, расположенных ниже.

По построению процесса на всяком шаге можно приближенно положить:

$$\gamma_1^{(N)} \approx \gamma_2^{(N)} \approx \dots \approx \gamma_n^{(N)} \approx \gamma^*. \quad (8.27)$$

Заметим, что в методе неопределенных множителей Лагранжа при допущении непрерывности x_i все частные производные (8.16) равны величине $\tilde{\lambda}$.

Если элементы системы достаточно высоконадежны, т. е. $\max_{1 \leq i \leq n} q_i \ll 1$, то практически при всех значениях x_i выполняется условие $Q_i(x_i + 1) \ll Q_i(x_i)$. В этом случае можно записать

$$\gamma_i(x_i) = \frac{Q_i(x_i) - Q_i(x_i + 1)}{c_i R_i(x_i)} \approx \frac{Q_i(x_i)}{c_i}. \quad (8.28)$$

Применив (8.28) к (8.27), получаем

$$c_i \gamma = Q_i(x_i), \quad (8.29)$$

что после суммирования по i дает

$$\gamma = Q^0 / \sum_{i=1}^n c_i. \quad (8.30)$$

В результате подстановки (8.30) в (8.29) имеем решение для i -й резервной группы в неявном виде:

$$Q_i(x_i) = Q^0 \left(c_i / \sum_{i=1}^n c_i \right). \quad (8.31)$$

Мы сформулировали правило для приближенного решения обратной задачи оптимального резервирования.

Если требования к надежности последовательной системы заданы в виде P^0 , то следует найти *квоту* ненадежности на каждый тип резервных элементов по формуле (8.31), а потом решить это уравнение относительно x_i .

ПРИМЕР 8.2. Рассмотрим ту же последовательную систему, что и в примере 8.1 (табл. 8.1). Требуется найти оптимальное число резервных элементов для системы, если время выполнения операции равно $t_0 = 10\,000$ ч, а требование по вероятности безотказной работы системы равно 0.9995. Резерв предполагается *холодным*.

Решение. Суммарная стоимость в четвертой колонке таблицы составляет 20 ст. ед. Находим сначала квоты ненадежности для элементов каждого типа по формуле (7.29):

$$Q_1 = 0.0005 \cdot 0.05 = 25 \cdot 10^{-6}; \quad Q_2 = 0.0005 \cdot 0.05 = 25 \cdot 10^{-6};$$

$$Q_3 = 0.0005 \cdot 0.4 = 200 \cdot 10^{-6}; \quad Q_4 = 0.0005 \cdot 0.4 = 200 \cdot 10^{-6};$$

$$Q_5 = 0.0005 \cdot 0.05 = 25 \cdot 10^{-6}; \quad Q_6 = 0.0005 \cdot 0.05 = 25 \cdot 10^{-6}.$$

При расчете надежности каждой группы с x_i запасными резервными элементами опять используем распределение Пуассона.

Находим для всех резервных групп величины $a = \lambda_i n_i t_0$:

$$a_1 = 1 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.05; \quad a_2 = 1 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.1;$$

$$a_3 = 1 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.05; \quad a_4 = 1 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.1;$$

$$a_5 = 8 \cdot 10^{-5} \cdot 5 \cdot 1000 = 0.4; \quad a_6 = 8 \cdot 10^{-5} \cdot 10 \cdot 1000 = 0.8.$$

В табл. 8.5 сведены те же значения величин $Q_i(x_i)$, что и в табл. 8.2. В каждом столбце серым цветом закрашена ячейка, где значение $Q_i(x_i)$ становится впервые меньше *квоты*, значения которых помещены в последней строке таблицы. Соответствующее решение (число необходимых резервных

Таблица 8.5

x_i	$Q_1(x_1)$	$Q_2(x_2)$	$Q_3(x_3)$	$Q_4(x_4)$	$Q_5(x_5)$	$Q_6(x_6)$
1	0.0476	0.0905	0.0476	0.0905	0.268	0.359
2	0.00119	0.00452	0.00119	0.00452	0.0536	0.144
3	1.98E-05	0.000151	1.98E-05	0.000151	0.00715	0.0383
4	2.48E-07	3.77E-06	2.48E-07	3.77E-06	0.000715	0.00767
5	2.48E-09	7.54E-08	2.48E-09	7.54E-08	5.72E-05	0.00123
6	2.06E-11	1.26E-09	2.06E-11	1.26E-09	3.81E-06	0.000164
7	1.47E-13	1.8E-11	1.47E-13	1.8E-11	2.18E-07	1.87E-05
8	9.22E-16	2.24E-13	9.22E-16	2.24E-13	1.09E-08	1.87E-06
9	5.12E-18	2.49E-15	5.12E-18	2.49E-15	4.84E-10	1.66E-07
10	2.56E-20	2.49E-17	2.56E-20	2.49E-17	1.94E-11	1.33E-08
Квота	25×10^{-6}	25×10^{-6}	200×10^{-6}	200×10^{-6}	25×10^{-6}	25×10^{-6}

Таблица 8.6

$C_{(1)} < C^0$	$C_{(2)} < C^0$	$Q_{(3)}^0 < Q_{(2)}^0$
	$C_{(2)} > C^0$	$Q_{(2)}^0 < Q_{(3)}^0 < Q_{(1)}^0$
$C_{(1)} > C^0$	$C_{(2)} < C^0$	$Q_{(3)}^0 < Q_{(2)}^0$
	$C_{(2)} > C^0$	$Q_{(2)}^0 < Q_{(3)}^0 < Q_{(1)}^0$

элементов) считывается в каждом столбце в закрашенных ячейках. Результат решения: $x_1 = 4$; $x_2 = 4$; $x_3 = 3$; $x_4 = 4$; $x_5 = 6$; $x_6 = 8$. Как нетрудно подсчитать по табл. 8.5, этому набору резервных элементов соответствует вероятность отказа 0.00018.

Полученное значение выше требуемого и при этом хуже того, что было получено методом покоординатного наискорейшего спуска. Это естественное следствие необходимости получения только целочисленных значений при фиксированных квотах на каждый тип элементов.

Решение прямой задачи оптимального резервирования удобнее всего проводить методом последовательных сравнений («методом вилки»). Выбираем некоторый уровень $Q_{(1)}^0$, находим для него решения $x_1^{(1)}, \dots, x_1^{(1)}$ и вычисляем соответствующее значение $C_{(1)}$. Возможны два варианта:

- (1) если $C_{(1)} < C^0$, то это означает, что имеются дополнительные ресурсы; выбираем новую квоту $Q_{(2)}^0$, $Q_{(2)}^0 < Q_{(1)}^0$, и продолжаем процедуру;
- (2) если $C_{(1)} > C^0$, то это означает перерасход отведенных ресурсов; выбираем новую квоту $Q_{(2)}^0$, $Q_{(2)}^0 < Q_{(1)}^0$, и продолжаем процедуру.

После второго шага мы имеем два значения $C_{(1)}$ и $C_{(2)}$. Если решение не получено, то делается следующий шаг, который выбирается так, как это пояснено в табл. 8.6. Такая процедура продолжается до тех пор, пока не будет найдено решение, удовлетворяющее заданным требованиям.

§ 8.4. МЕТОД ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

8.4.1. Описание метода

Суть метода динамического программирования* заключается в пошаговой процедуре отыскания условных оптимальных решений, которая приводит на последнем шаге к нахождению абсолютного оптимума.

* Метод разработан американским математиком Р. Беллманом (Bellman R. E., 1920—1984), поэтому часто именуется методом Беллмана.

Целевые функции в задачах оптимального резервирования обычно являются унимодальными.

Формально прямая задача оптимального резервирования записывается в виде функционального уравнения

$$f_n(R^0) = \min_{\substack{R_n(x) \geq R^0 \\ x_n = 0, 1, 2, \dots}} \left\{ c_n x_n + f_{n-1} \left(\frac{R^0}{R_n(x_n)} \right) \right\}. \quad (8.32)$$

В правой части под знаком min стоят две функции, одна из них, $c_n x_n$, известна. Вторая функция дает оптимальное решение для остальных $(n - 1)$ подсистем при условии, что на n -ю подсистему затрачено $c_n x_n$ ст. ед. Если бы была известна и функция f_{n-1} , то решение было бы очевидно.

Однако функция f_{n-1} неизвестна, ее сначала нужно построить на основании функции f_{n-2} , которая тоже неизвестна! Поэтому необходимо построить систему рекуррентных уравнений и затем начать численное решение с конца.

Аналогичная картина складывается и для обратной задачи оптимального резервирования. Здесь необходимо решить функциональное уравнение типа

$$f_n(C^0) = \max_{\substack{0 \leq c_n x_n \leq C^0 \\ x_n = 0, 1, 2, \dots}} \{ R_n(x_n) \cdot f_{n-1}(C^0 - c_n x_n) \}. \quad (8.33)$$

В правой части под знаком max стоят две функции, одна из них, $R_n(x_n)$, известна. Вторая функция дает оптимальное решение для остальных $(n - 1)$ подсистем при условии, что на них осталась возможность затратить $(C^0 - c_n x_n)$ ст. ед.

Метод динамического программирования является по существу хорошо организованным перебором с поэтапным отсеком неперспективных вариантов.

8.4.2. Модифицированный метод динамического программирования

Пояснить процедуру динамического программирования проще и удобнее на его модификации*, алгоритме Кеттелля. Суть этого метода такова. Рассмотрим две какие-либо подсистемы, например n -ю и $(n - 1)$ -ю. Составим предварительно для каждой из них таблицу вида табл. 8.7.

* Метод разработан американским математиком Дж. Кеттеллем.

Таблица 8.7

x_i	0	1	2	3	...
$R_i(x_i)$	$R_i(0)$	$R_i(1)$	$R_i(2)$	$R_i(3)$...
$C_i(x_i)$	0	c_i	$2c_i$	$3c_i$...

Далее построим табл. 8.8, в которую сведем все возможные парные комбинации этих двух систем; в ней можно считать, что $R(K, N) = R_n(K) \cdot R_{n-1}(N)$ при условии независимости отказов подсистем.

Из табл. 8.8, находят доминирующую последовательность, которой принадлежат все ее ячейки, удовлетворяющие условиям следующей процедуры.

- Пусть некоторая ячейка включена в доминирующую последовательность под номером N . Отыскивается следующая ячейка, у которой значение стоимости выше.
- Если при этом ее показатель надежности ниже, чем в N -й ячейке, то она исключается из дальнейшего рассмотрения, и процедура продолжается.
- Если же показатель надежности в следующей ячейке выше, чем в N -й ячейке, то эта новая ячейка включается в доминирующую последовательность.

Таблица 8.8

$n-1$	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$...
0	0 $R(0,0)$	c_n $R(0,1)$	$2c_n$ $R(0,2)$	$3c_n$ $R(0,3)$	$4c_n$ $R(0,4)$...
1	c_{n-1} $R(1,0)$	$c_{n-1} + c_n$ $R(1,1)$	$c_{n-1} + 2c_n$ $R(1,2)$	$c_{n-1} + 3c_n$ $R(1,3)$	$c_{n-1} + 4c_n$ $R(1,4)$...
2	$2c_{n-1}$ $R(2,0)$	$2c_{n-1} + c_n$ $R(2,1)$	$2c_{n-1} + 2c_n$ $R(2,2)$	$2c_{n-1} + 3c_n$ $R(2,3)$	$2c_{n-1} + 4c_n$ $R(2,4)$...
3	$3c_{n-1}$ $R(3,0)$	$3c_{n-1} + c_n$ $R(3,1)$	$3c_{n-1} + 2c_n$ $R(3,2)$	$3c_{n-1} + 3c_n$ $R(3,3)$	$3c_{n-1} + 4c_n$ $R(3,4)$...
4	$4c_{n-1}$ $R(4,0)$	$4c_{n-1} + c_n$ $R(4,1)$	$4c_{n-1} + 2c_n$ $R(4,2)$	$4c_{n-1} + 3c_n$ $R(4,3)$	$4c_{n-1} + 4c_n$ $R(4,4)$...
...

Иными словами, если разместить все пары (C, R) , помещенные в ячейках табл. 8.8, на плоскости (рис. 8.5), то члены доминирующей последовательности образуют монотонную оболочку.

Итак, получена первая доминирующая последовательность для двух подсистем

$$\{(c_1^*, R_1^*; Y_1^{(1)}), (c_2^*, R_2^*; Y_2^{(1)}), \dots, (c_k^*, R_k^*; Y_k^{(1)}), \dots\}; \quad (8.34)$$

здесь $Y_k^{(1)}$ — вектор, в который входит число элементов n -го и $(n-1)$ -го типов для k -го члена доминирующей последовательности на 1-м шаге процесса. Иными словами, вектор $Y_k^{(1)}$ содержит координаты соответствующей ячейки табл. 8.8.

К вновь полученной доминирующей последовательности применим опять прежнюю процедуру, используя на следующем шаге $(n-2)$ -ю подсистему, и составим табл. 8.8. Здесь в заголовках колонок стоят номера, соответствующие порядковым номерам членов доминирующей последовательности из табл. 8.8. Эти номера помечены звездочкой.

К табл. 8.9 применяется вновь процедура построения доминирующей последовательности. На последнем шаге рассмотренного

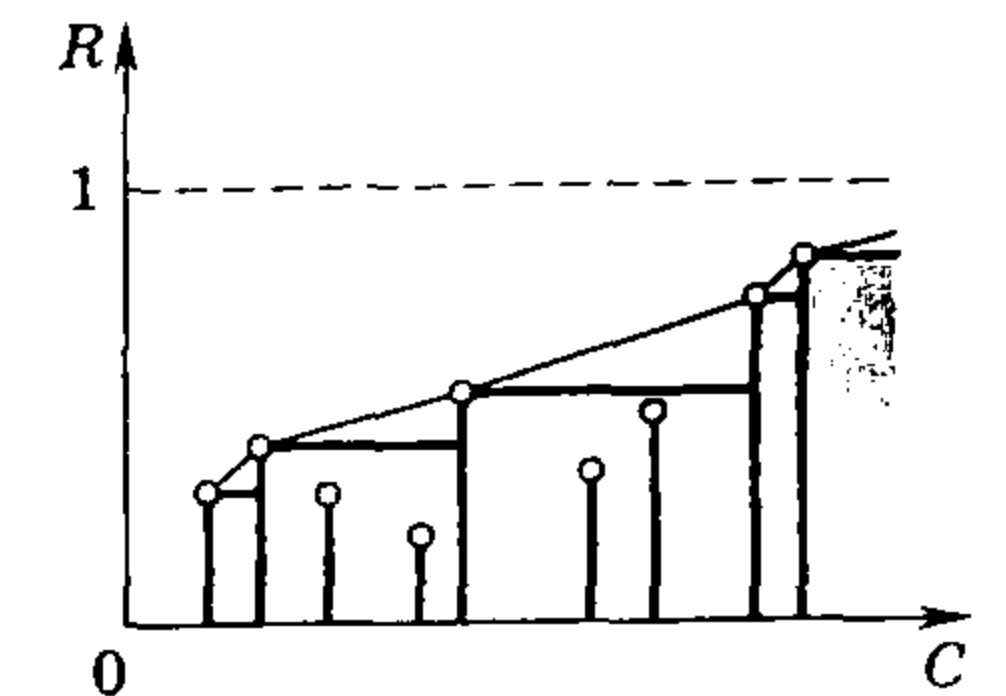


Рис. 8.5

Таблица 8.9

$n-2$	1*	2*	3*	4*	5*	...
0	c_1^* $R(0, 1^*)$	c_2^* $R(0, 2^*)$	c_3^* $R(0, 3^*)$	c_4^* $R(0, 4^*)$	c_5^* $R(0, 5^*)$...
1	$c_1^* + c_{n-2}$ $R(1, 1^*)$	$c_2^* + c_{n-2}$ $R(1, 2^*)$	$c_3^* + c_{n-2}$ $R(1, 3^*)$	$c_4^* + c_{n-2}$ $R(1, 4^*)$	$c_5^* + c_{n-2}$ $R(1, 4)$...
2	$c_1^* + 2c_{n-2}$ $R(2, 1^*)$	$c_2^* + 2c_{n-2}$ $R(2, 2^*)$	$c_3^* + 2c_{n-2}$ $R(2, 3^*)$	$c_4^* + c_{n-2}$ $R(2, 4^*)$	$2c_5^* + 2c_{n-2}$ $R(2, 4)$...
3	$c_1^* + 3c_{n-2}$ $R(3, 1^*)$	$c_2^* + 3c_{n-2}$ $R(3, 2^*)$	$c_3^* + 3c_{n-2}$ $R(3, 3^*)$	$c_4^* + 3c_{n-2}$ $R(3, 4^*)$	$c_5^* + 3c_{n-2}$ $R(3, 4)$...
4	$c_1^* + 4c_{n-2}$ $R(4, 1^*)$	$c_2^* + 4c_{n-2}$ $R(4, 2^*)$	$c_3^* + 4c_{n-2}$ $R(4, 3^*)$	$c_4^* + 4c_{n-2}$ $R(4, 4^*)$	$2c_5^* + 4c_{n-2}$ $R(4, 4)$...
...

процесса оказывается построенной окончательная доминирующая последовательность

$$(\tilde{C}_1, \tilde{R}_1, Y_1^{(n)}), (\tilde{C}_2, \tilde{R}_2, Y_2^{(n)}), (\tilde{C}_3, \tilde{R}_3, Y_3^{(n)}), \dots, (\tilde{C}_k, \tilde{R}_k, Y_k^{(n)}), \dots, \quad (8.35)$$

где в $Y_k^{(n)}$ содержится вся информация о том, сколько элементов каждого типа использовано в данном варианте. Действительно, $Y_k^{(n)}$ содержит $x_k^{(n)}$ и соответствующий вектор $Y_k^{(n-1)}$; $Y_k^{(n-1)}$ содержит $x_k^{(n-1)}$ и соответствующий вектор $Y_k^{(n-2)}$, и т. д.

Доминирующая последовательность представляет собой совокупность *неулучшаемых решений* (множество Парето).

Для решения прямой задачи оптимального резервирования (минимизации стоимости при заданном показателе надежности) выбирается член k -й доминирующей последовательности такой, что $\tilde{R}_k \geq R^0$, а $\tilde{R}_{k-1} < R^0$, при этом стоимость системы будет \tilde{C}_k , а число элементов каждого типа содержится в векторе $Y_k^{(n)}$.

Решение обратной задачи оптимального резервирования (максимизация надежности при ограничении на суммарную стоимость резервных элементов) сводится к выбору k -го члена доминирующей последовательности такого, что $\tilde{C}_k \leq C^0$, а $\tilde{C}_{k+1} > C^0$, при этом надежность такой системы будет \tilde{R}_k , а число элементов каждого типа содержится в векторе $Y_k^{(n)}$.

Процедуру лучше всего пояснить с помощью конкретного числового примера.

ПРИМЕР 8.3. Рассмотрим последовательную систему, состоящую из четырех резервных групп (подсистем) с *холодным* резервом. Характеристики элементов системы приведены в табл. 8.10.

Требуется найти оптимальное число резервных элементов для системы, если время выполнения операции равно $t_0 = 1000$ ч, для двух случаев:

- (1) требование по вероятности безотказной работы системы равно 0,999;
- (2) допустимые затраты на резервные элементы составляют 30 ст. ед.

Решение. Составим таблицу, содержащую значения $Q_i(x_i)$ и $c_i(x)$ для каждой резервной группы (табл. 8.11). Затем построим доминирующие последовательности, необходимые для решения.

Имеются два способа последовательного построения доминирующих последовательностей. Первый из них описан выше и представлен для рассматриваемого в примере случая на рис. 8.6, а. Второй способ — дихотомическое «сворачивание» подсистем, представленное на рис. 8.6, б.

Таблица 8.10

Номер элемента	Число элементов, n_k	Интенсивность отказов, λ_k (10^{-5} 1/ч)	Стоимость элемента, c_k
1	5	1	8
2	10	1	8
3	5	8	1
4	10	8	1

Таблица 8.11

x_i	1-я группа	2-я группа	3-я группа	4-я группа
0	0.0476	0.0905	0.268	0.359
1	0.00119	0.00452	0.0536	0.144
2	1.98E-05	0.000151	0.00715	0.0383
3	2.48E-07	3.77E-06	0.000715	0.00767
4	2.48E-09	7.54E-08	5.72E-05	0.00123
5	2.06E-11	1.26E-09	3.81E-06	0.000164
6	1.47E-13	1.8E-11	2.18E-07	1.87E-05
7	9.22E-16	2.24E-13	1.09E-08	1.87E-06
8	5.12E-18	2.49E-15	4.84E-10	1.66E-07
9	2.56E-20	2.49E-17	1.94E-11	1.33E-08

* В каждой ячейке таблицы верхнее полужирное число соответствует стоимости, нижнее, светлое, — надежности.

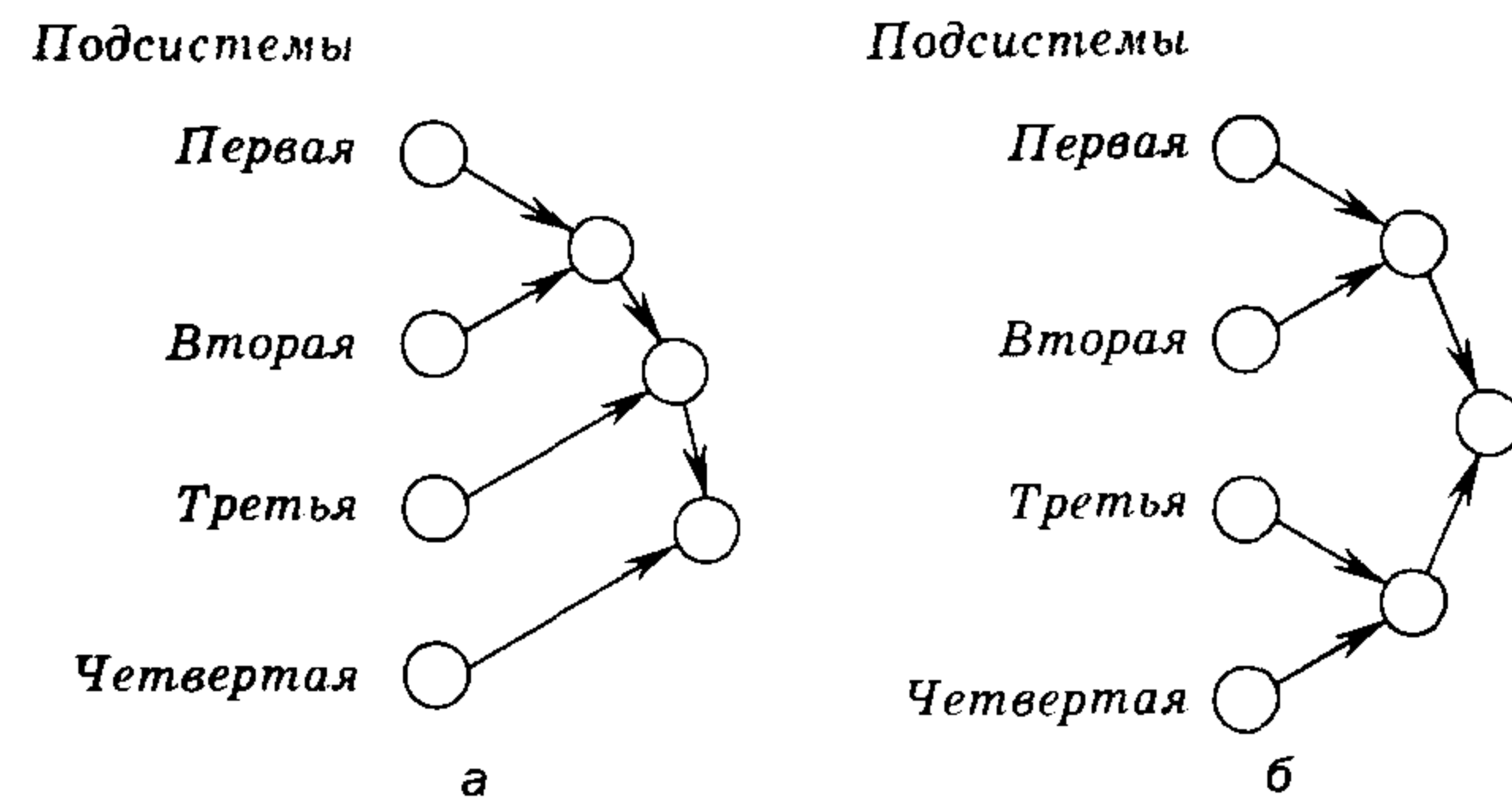


Рис. 8.6

Заметим, что на практике, особенно при ручном счете, второй способ оказывается предпочтительнее. Мы применим его при решении данного примера. Итак, построим сначала доминирующую последовательность для 1-й и 2-й подсистем и сведем ее в табл. 8.12. Здесь приведена начальная часть доминирующей последовательности, образованной сверткой 1-й и 2-й подсистем. В кружочках проставлены номера соответствующих членов доминирующей последовательности. Будем в дальнейшем обозначать через $y_j^{(1)}$ вектор, составленный из соответствующего числа элементов 1-го и 2-го типов: так, например, $y_{(3)} = (x_1 = 1, x_2 = 2)$.

Построим теперь аналогичную таблицу для 3-й и 4-й подсистем — табл. 8.13.

В ней приведена начальная часть доминирующей последовательности, образованной сверткой 3-й и 4-й подсистем. В кружочках проставлены номера

Таблица 8.12

	x_2	0	1	2	3	4	...
x_1	2-я 1-я	0 0.0905	1 0.00452	2 0.000151	3 3.77E-06	4 7.54E-08	...
0	0.0476	0.1381	① 0.05212	16 0.047751	24 0.047604	32 0.0476	...
1	0.00119	0.09169	② 0.00571	③ 24 0.001341	32 0.001194	40 0.00119	...
2	1.98E-05	0.09052	24 0.00454	④ 32 0.000171	⑤ 40 2.36E-05	48 1.99E-05	...
3	2.48E-07	0.0905	32 0.00452	40 0.000151	⑥ 48 4.02E-06	⑦ 56 3.23E-07	...
...

Таблица 8.13

	x_4	0	1	2	3	4	5	6	7	...
x_3	4-я 3-я	0.359	0.144	0.0383	0.0383	0.00123	1.64E-04	1.87E-05	1.87E-06	...
0	0.268	0.627	① 0.412	0.3063	0.3063	0.26923	0.268164	0.268019	0.268002	...
1	0.0536	0.4126	② 0.1976	③ 0.0919	0.0919	0.05483	0.053764	0.053619	0.053602	...
2	0.00715	0.36615	0.15115	④ 0.04545	0.04545	⑥ 0.00838	0.007314	0.007169	0.007152	...
3	0.000715	0.359715	0.144715	⑤ 0.039015	0.039015	⑦ 0.001945	⑧ 0.000879	0.000734	0.000717	...
4	5.72E-05	3.59E-01	1.44E-01	⑥ 3.84E-02	3.84E-02	8 1.29E-03	⑨ 2.21E-04	⑩ 7.59E-05	11 5.91E-05	...
5	3.81E-06	3.59E-01	1.44E-01	⑦ 3.83E-02	3.83E-02	9 1.23E-03	10 1.68E-04	⑪ 2.25E-05	⑫ 5.68E-06	...
6	2.18E-07	3.59E-01	1.44E-01	⑧ 3.83E-02	3.83E-02	10 1.23E-03	11 1.64E-04	12 1.89E-05	13 2.09E-06	...

соответствующих членов доминирующей последовательности. Аналогично предыдущему обозначим через $y_j^{(2)}$ вектор, составленный из соответствующего числа элементов 1-го и 2-го типов: так, например, $y_6^{(2)} = (x_1 = 2, x_2 = 4)$.

Таблица 8.14

		j					
		0	1	2	3	4	
$y_k^{(2)}$	1-я	0	8	16	24	32	
	2-я	0.1381	0.0521	0.00571	0.00134	0.000171	
k	0	0	8	16	24	32	
	0	0.627	0.678511	0.646433	0.62913	0.62834	0.627171
	1	1	① 1	9	17	25	33
	1	0.412	0.493203	0.442635	0.41771	0.41334	0.412171
	2	2	② 2	10	18	26	34
	2	0.1976	0.3357	0.239405	0.20331	0.19894	0.197771
	3	3	③ 3	11	19	27	35
	3	0.0919	0.23	0.139212	0.09761	0.09324	0.092071
	4	4	④ 4	⑫ 12	⑳ 20	㉘ 28	36
	4	0.0455	0.1836	0.0976	0.05121	0.04684	0.045671
	5	5	⑤ 5	⑬ 13	㉑ 21	㉙ 29	37
	5	0.039	0.1771	0.0911	0.04471	0.04034	0.039171
6	6	⑥ 6	⑭ 14	㉒ 22	㉚ 30	38	
6	0.00838	0.14648	0.06048	0.01409	0.00972	0.008551	
7	7	⑦ 7	⑮ 15	㉓ 23	㉛ 31	39	
7	0.001945	0.140045	0.054045	0.007655	0.003285	0.002116	
8	8	⑧ 8	⑯ 16	㉔ 24	㉜ 32	㉟ 40	
8	0.000879	0.138979	0.052979	0.006589	0.002219	0.00105	
9	9	⑨ 9	⑰ 17	㉕ 25	㉝ 33	41	
9	0.000221	0.138321	0.052321	0.005931	0.001561	0.000392	
10	10	⑩ 10	⑱ 18	㉖ 26	㉞ 34	42	
10	7.59E-05	0.138176	0.052176	0.005786	0.001416	0.000247	
11	11	⑪ 11	⑲ 19	㉗ 27	㉟ 35	43	
11	2.25E-05	0.138123	0.052123	0.005733	0.001363	0.000194	
12	12	12	20	28	㉠ 36	44	
12	5.68E-06	0.138106	0.052106	0.005716	0.001346	0.000177	

Теперь перейдем к построению завершающей доминирующей последовательности, используя табл. 8.12 и табл. 8.13. Результаты сведены в табл. 8.14.

Решение прямой задачи (1) достигается на 37-м члене финальной доминирующей последовательности в табл. 8.14 при стоимости системы, равной 40 ст. ед.*.

Заметим, что это решение соответствует значениям $y_5^{(1)}$ (пятый столбец) и $y_8^{(2)}$ (восьмая строка).

По табл. 8.12 находим, что $y_5^{(1)} = (x_1 = 2, x_2 = 3)$, и по табл. 8.13, что $y_8^{(2)} = (x_1 = 3, x_2 = 5)$.

Решение обратной задачи (2) достигается на 30-м члене финальной доминирующей последовательности при надежности системы, равной $1 - 0.00972 = 0.99028 \approx 0.99$. Решение соответствует значениям $y_3^{(1)}$ и $y_6^{(2)}$. По табл. 8.12 находим, что $y_3^{(1)} = (x_1 = 2, x_2 = 2)$, и по табл. 8.13, что $y_6^{(2)} = (x_1 = 2, x_2 = 4)$.

§ 8.5. ОПТИМАЛЬНОЕ РЕЗЕРВИРОВАНИЕ ПРИ НЕСКОЛЬКИХ ОГРАНИЧЕНИЯХ

Подобного рода задачи встречаются на практике при проектировании различного рода мобильных объектов (самолетов, кораблей, передвижной техники), когда в расчет приходится брать не только ограничения по суммарной стоимости резервных элементов, но также и другие параметры, например вес и габариты (т. е. объем) резервных частей.

Чаще всего такие параметры системы являются противоречивыми: более легкие и менее объемные элементы стоят дороже. Точное решение задач весьма трудоемко. В связи с этим привлекает внимание метод *приведенной стоимости*. Будем говорить о *ресурсе-1* и *ресурсе-2* вместо стоимости и веса. Рассмотрим случай двух типов ресурсов, каждый со своим ограничением. Суть метода заключается в следующем: вместо двумерной задачи оптимального резервирования решаются несколько одномерных задач с приведенной стоимостью. *Весовые* коэффициенты $p^{(k)}$, $0 \leq p^{(k)} \leq 1$ определенным образом меняются, и при каждом их значении решается одномерная задача для приведенной стоимости:

$$c_j^{(k)} = p^{(k)}c_{1j} + (1 - p^{(k)})c_{2j},$$

где c_{1j} — затраты ресурса-1 на j -й резервный элемент, c_{2j} — затраты ресурса-2 на него же, $c_j^{(k)}$ — приведенный ресурс j -го ре-

* Пренебрежем незначительным превышением показателя ненадежности над допустимым значением, составляющим 0.001.

зервного элемента, взвешенный по $p^{(k)}$, а $p^{(k)}$ — весовой коэффициент, пробегающий некоторые дискретные значения в интервале $[0,1]$. В процессе решения для каждого значения приведенного ресурса проводятся контрольные проверки, не нарушено ли хотя бы одно из ограничений на ресурс-1 или ресурс-2. После решения задач с различными приведенными стоимостями выбирается лучшее из них, удовлетворяющее всем требуемым ограничениям.

Отметим, что во многих практических задачах оказывается целесообразным решить в двумерном случае две пограничные задачи с $p^{(k)} = 0$ и $p^{(k)} = 1$. Это означает, что на практике достаточно хорошее решение может получиться, если решить n одномерных задач, используя для процедуры наискорейшего спуска один из параметров, а остальные только контролировать, чтобы не нарушались заданные ограничения на них. В качестве решения при этом выбирается то, которое обеспечивает максимальное значение показателя надежности системы.

В заключение отметим, что оптимальное резервирование — важная практическая проблема. Особую роль она играет при рассмотрении оптимального обеспечения систем запасными элементами.

В качестве приемлемого метода решения практически всех инженерных задач можно рекомендовать метод покоординатного наискорейшего спуска, наиболее простой из всех рассмотренных выше.

Это является оправданным и с математической точки зрения: в инженерной практике статистическая информация об элементах настолько недостоверна, что использование строгих методов оптимизации представляется некорректным.

Глава 9

Обеспечение систем типовыми элементами замены

Одной из важных задач технического обслуживания больших территориальных систем является организация иерархической системы складов запасных элементов, необходимых для поддержания непрерывного функционирования обслуживаемых объектов. Рассматривается модель такой системы складов, их оптимального снабжения и размещения в ней оптимального числа запасных элементов.

§ 9.1. ОБЩЕЕ ОПИСАНИЕ СИСТЕМЫ

Многие современные технические системы представляют собой большие территориальные комплексы с размещением объектов (подсистем) на огромной территории. Успешное функционирование таких систем возможно лишь при соответствующей организации их технического обслуживания. Одной из важнейших задач технического обслуживания является своевременное и надежное снабжение систем типовыми элементами замены (ТЭЗ), необходимыми для замены отказавших элементов в объектах системы. При этом сами объекты, являясь в некотором смысле однотипными (т. е. имеющими аналогичную структуру, предназначенными для выполнения одинаковых функций, и т. п.), могут различаться по числу входящих в них элементов различного типа. Примером могут служить наземные базовые станции спутниковых систем телекоммуникации.

Каждый объект большой территориально-распределенной системы должен иметь свой собственный, индивидуальный запас ТЭЗ. Однако возникают вопросы: на какой период времени нужно обеспечивать объект ТЭЗ? Как пополнять запасы ТЭЗ при их исчерпании? Какова допустимая вероятность того, что не найдется замены отказавшему элементу?

Подчеркнем, что речь идет не о регулярном снабжении запасными элементами и связанными с этим затратами, а о создании

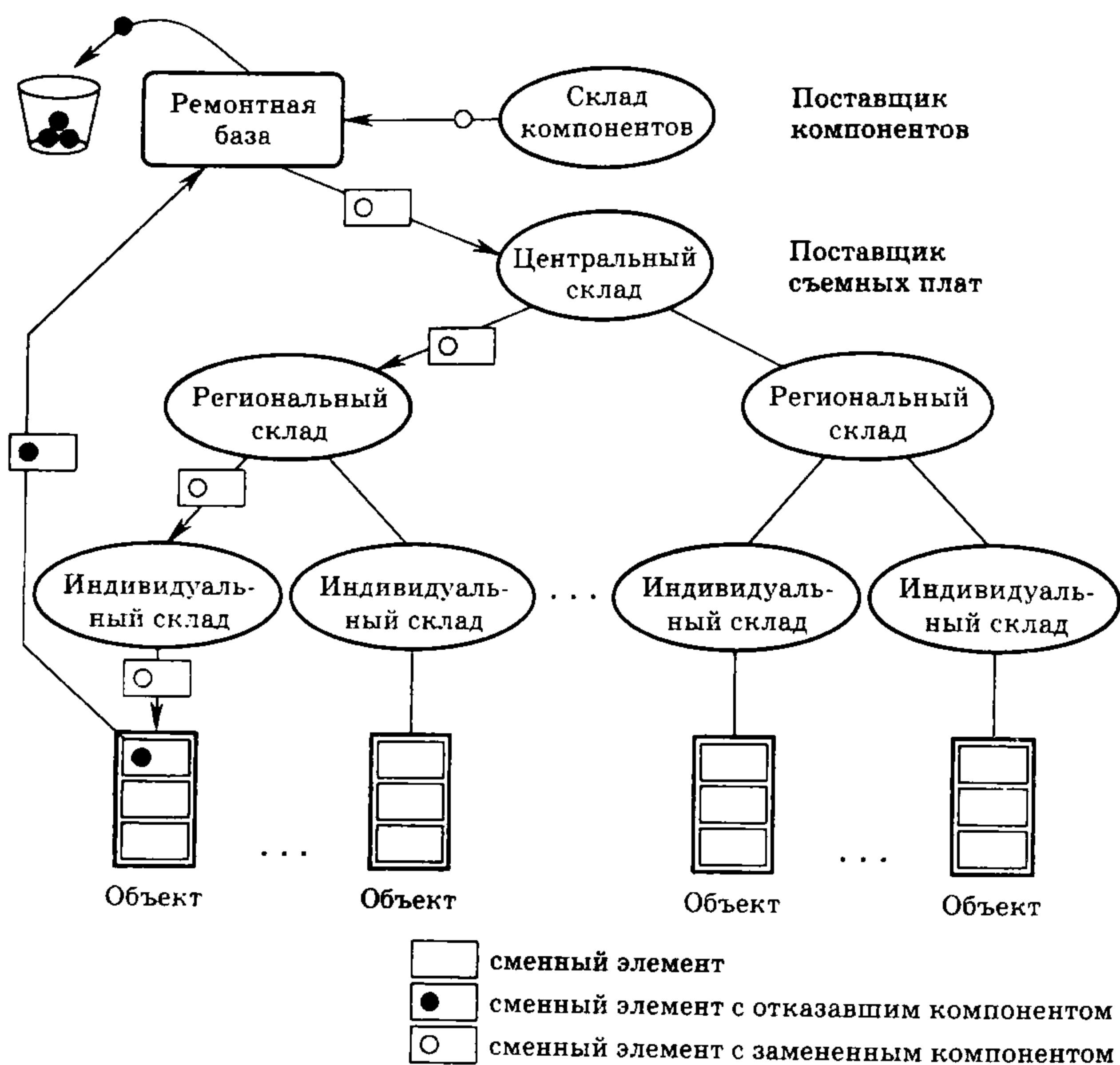


Рис. 9.1

одноразового гарантийного запаса необходимыми элементами, который служит для безотказного *выживания* системы в течение времени доставки заказанных запчастей.

Поскольку индивидуальный склад сам по себе является частью разветвленной складской системы, возникают и другие вопросы. Например, те же наземные станции спутниковой системы телекоммуникации разбросаны по всему свету, поэтому возникает дополнительная задача: как обеспечить своевременную доставку ТЭЗ по мере их исчерпания на индивидуальном складе? Обычно оказывается целесообразным использовать региональные склады, являющиеся промежуточным звеном снабжения; сами они пополняются периодически из центрального склада. Центральный склад связан с ремонтной базой, на которой производится восстановление ТЭЗ, отказавших на объектах системы. Обычно центральный склад и ремонтная база размещаются территориально в одном и том же месте.

Блок-схема, поясняющая процесс технического обслуживания, приведена на рис. 9.1. Имеется в виду, что ТЭЗ состоит из невосстанавливаемых компонентов, отказ которых требует их замены на ремонтной базе.

Система технического обслуживания должна рассматриваться как неотъемлемая часть самой технической системы, которую она обслуживает. Проектирование системы обслуживания такой системы завершено к моменту завершения проектирования такой системы.

Требования к системе обслуживания определяются требованиями к самой системе: чем надежнее должна быть проектируемая система, тем жестче требования к качеству и надежности обслуживания.

Итак, типовая система обслуживания большой территориально-распределенной системой включает в свой состав:

- ✓ центральный склад ТЭЗ;
- ✓ региональные склады ТЭЗ;
- ✓ индивидуальные склады ТЭЗ;
- ✓ ремонтную базу для восстановления работоспособности ТЭЗ путем замены в них отказавших компонентов;
- ✓ склада компонентов, необходимых для ремонта отказавших ТЭЗ;
- ✓ поставщиков компонентов, необходимых для ремонта, а также пополнения общего запаса ТЭЗ, которые могут отказаться до неремонтируемого уровня.

Иерархическая система складов, которая является основным предметом рассмотрения в этой главе, представлена на рис. 9.2.

Анализ и проектирование таких систем обслуживания представляет собой сложную инженерно-экономическую задачу.

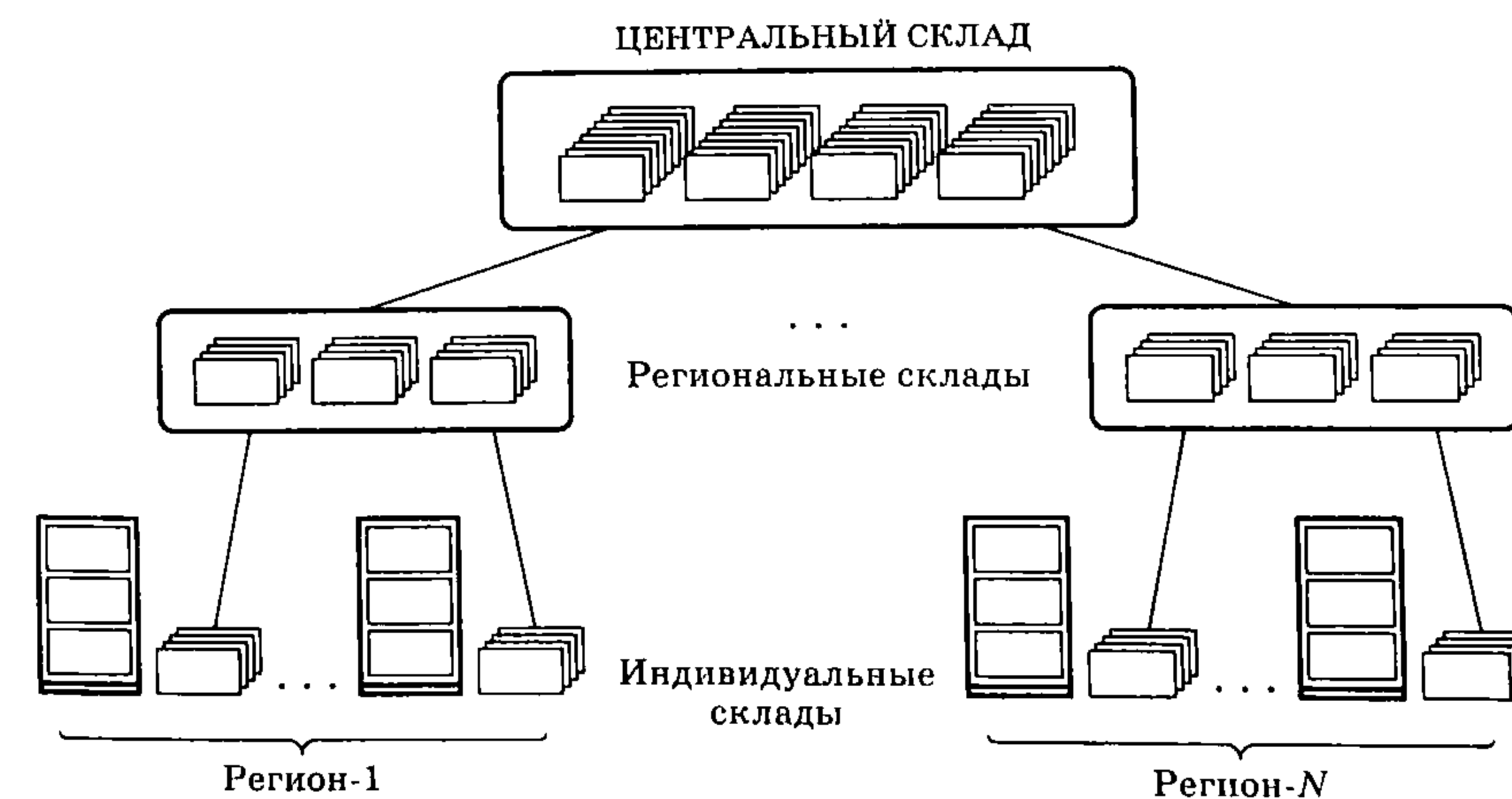


Рис. 9.2

§ 9.2. АНАЛИЗ ИНДИВИДУАЛЬНОГО СКЛАДА

ОБЩЕЕ ОПИСАНИЕ. Пусть объект включает в себя N различных типов элементов, причем число элементов k -го типа равно n_k . Каждый элемент k -го типа характеризуется интенсивностью отказов λ_k . Индивидуальный склад обычно после каждого очередного отказа объекта посылает заявку на доставку соответствующего исправного ТЭЗ с регионального склада. Пусть время такой доставки равно Δ . Обычно длительность этого интервала составляет от пяти до десяти дней, что соответствует времени не экстренной почтовой доставки. Фактически индивидуальный склад играет роль простого буфера: он позволяет обслуживаемому объекту *выжить* в течение времени доставки ТЭЗ с регионального склада.

Индивидуальный склад содержит x_1, x_2, \dots, x_n ТЭЗ различных типов. После очередной замены элемента k -го типа на складе остается $(x_k - 1)$ ТЭЗ этого типа. Отказ обслуживаемого объекта произойдет, если на интервале доставки Δ произойдет x_k отказов элементов именно k -го типа, так как последний отказ уже не будет обеспечен запасным ТЭЗ данного типа.

Задача заключается в том, чтобы найти оптимальное число ТЭЗ каждого типа x_1, x_2, \dots, x_n , чтобы обеспечивались заданные требования на надежность индивидуального склада при минимальных затратах. Для решения этой задачи необходимо иметь формулу коэффициента готовности индивидуального склада — его основной характеристики.

ПРАВИЛО ПОПОЛНЕНИЯ ИНДИВИДУАЛЬНОГО СКЛАДА. Индивидуальный склад пополняется по заявке, посылаемой в региональный склад в момент отказа элемента какого-либо типа на том объекте, который обслуживается данным складом. Такой способ минимизирует число ТЭЗ на складе, однако приводит к возрастанию затрат на доставку.

На рис. 9.3 представлен случай когда индивидуальный склад отказывает, имея два ТЭЗ k -го типа: первый из них расходуется в момент отказа, тогда же посылается заявка в региональный центр; второй служит для следующей замены, когда в системе появляется еще один отказ элемента того же типа. Следующий отказ элемента k -го типа приводит к отказу индивидуального склада.

Для каждого типа ТЭЗ картина замены имеет такой же вид.

В случае индивидуальных складов важную роль играет учет структуры обслуживаемого объекта (внутренний резерв). Отказ очередного элемента может оказаться не критическим, если име-

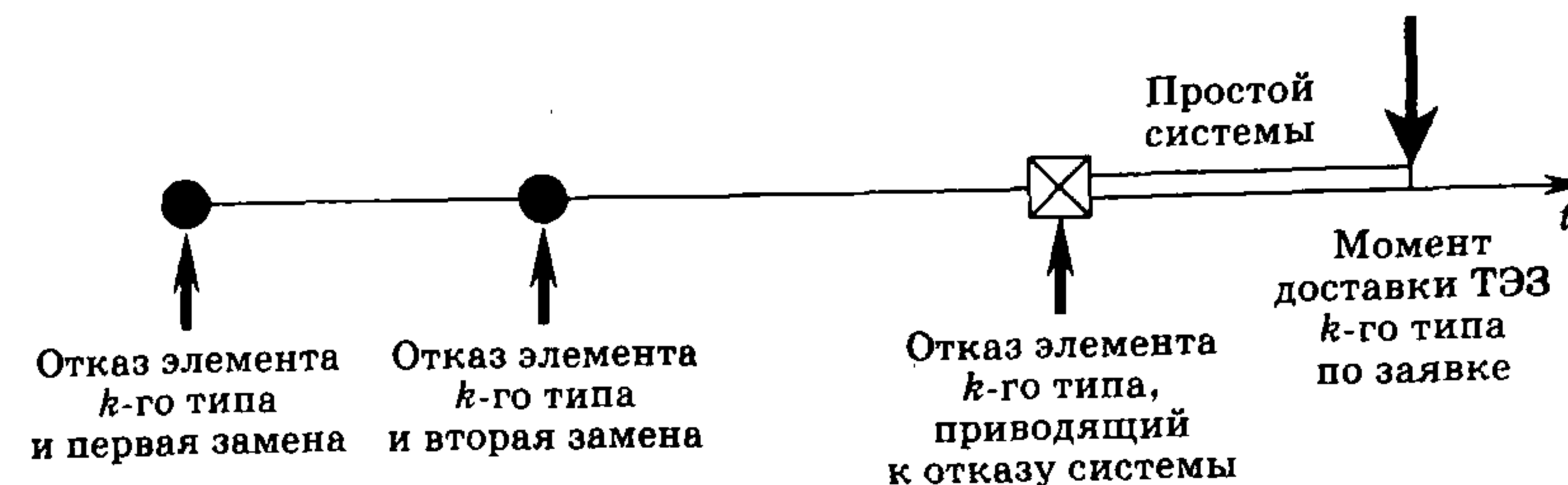


Рис. 9.3

ется *горячий* резерв. Для простоты описания опустим подобную возможность, т. е. будем считать, что склад должен обеспечить комплектность штатной структуры системы. Учет резерва не представляет принципиальных трудностей.

ТРЕБОВАНИЯ К ИНДИВИДУАЛЬНОМУ СКЛАДУ. Индивидуальный склад сменных ТЭЗ необходим для практически мгновенной замены отказавшего элемента на работоспособный. Чтобы решить задачу нахождения оптимального числа запасных ТЭЗ каждого типа, необходимо сформулировать требования к надежности индивидуального склада. Чаще всего эти требования базируются на предыдущем практическом опыте проектирования и эксплуатации аналогичных складских систем. Однако можно указать и на более обоснованный способ: задание требований на надежность складской системы, базирующееся на требованиях по надежности к самому обслуживаемому объекту.

Пусть известно, что требуемое значение коэффициента готовности обслуживаемого объекта равно A . При проектировании самого объекта обычно предполагают, что в запасных элементах нехватки не будет. Иначе говоря, математическая модель объекта учитывает время на замену, время на переключения с основного элемента на резервный и т. п., но возможные простои объекта, вызванные нехваткой запасных ТЭЗ, не учитываются.

Возникает вопрос: какую степень ухудшения коэффициента готовности мы готовы допустить за счет неидеальности системы запасов? Предположим, что $A_{\text{объект}} = 0.995$, что эквивалентно простоям в среднем пяти часов на каждую тысячу часов работы. Допустимо ли, чтобы вместо пяти часов в среднем объект простаивал, скажем, шесть часов? Какое увеличение коэффициента простоя $U_{\text{сист}}$, $U_{\text{сист}} = 1 - A_{\text{сист}}$ принять: на 20 или на 1%? Это является прерогативой лица, принимающего решение. Когда такое решение принято и допустимое значение КГ определено, можно использовать следующую процедуру.

Обозначим через β ту долю $U_{\text{сист}}$, снижение на которую считается допустимым:

$$U_{\text{инд}} = \beta U_{\text{сист}}. \quad (9.1)$$

Величина β выбирается в диапазоне от 1% до 10%, осуществляется на основании анализа того, что является более эффективным: увеличивать надежность обслуживаемого объекта или индивидуального склада.

Интенсивность отказов индивидуального склада. После отказа элемента k -го типа за время доставки ТЭЗ с регионального склада на обслуживаемом объекте не должно отказать более $(x_k - 1)$ элементов того же типа. Вероятность этого события равна*

$$r_k(x_k) = \sum_{j=0}^{x_k-1} \frac{(\lambda_k n_k \Delta)^j}{j!} \cdot \exp(-\lambda_k n_k \Delta), \quad (9.2)$$

где λ_k — интенсивность отказа элемента k -го типа, n_k — число этих элементов в обслуживаемой системе, а Δ — длительность доставки по запросу.

Для расчетов удобнее использовать дополнительную вероятность, т. е. вероятность отказа системы за время доставки ТЭЗ k -го типа $q_k(x_k)$, которая имеет следующий вид:

$$q_k(x_k) = \sum_{j=x_k}^{\infty} \frac{(\lambda_k n_k \Delta)^j}{j!} \cdot \exp(-\lambda_k n_k \Delta). \quad (9.3)$$

При $q_k(x_k) \ll 1/n$, т. е. в случае высоконадежной складской системы, эта вероятность может быть записана приближенно как

$$q_k(x_k) \approx \frac{(\lambda_k n_k \Delta)^{x_k}}{x_k!}. \quad (9.4)$$

Таким образом, каждый отказ элемента k -го типа может привести к отказу объекта с вероятностью $q_k(x_k)$. Поток отказов элементов k -го типа является пуассоновским с параметром $\lambda_k n_k$, поэтому поток отказов у системы из-за этих отказов в соответствии с теоремой о разрежении потоков будет равен

$$\Lambda_k = \lambda_k n_k q_k(x_k) \approx \lambda_k n_k \frac{(\lambda_k n_k \Delta)^{x_k}}{x_k!}. \quad (9.5)$$

Результирующий поток отказов обслуживаемого объекта будет равен

$$\Lambda_{\text{инд}} = \sum_{k=1}^N \Lambda_k, \quad (9.6)$$

* Предполагается, что поток отказов элементов каждого типа является пуассоновским. Обоснованность такого рода предположения опирается на теорему о суперпозиции потоков отказов.

где поток Λ_k определен выражением (9.5), а N — число различных типов элементов.

Среднее время простоя индивидуального склада. Рассмотрим ситуацию, когда обслуживаемый объект отказал из-за отказа элементов k -го типа. Это может произойти, если число отказов, возникших за время доставки Δ , будет равно или больше x_k , т. е. превысит число запасных ТЭЗ этого типа на индивидуальном складе. Рассматривая высоконадежный объект, мы можем предположить, что если он отказал, то в интервале времени Δ возникло x_k отказов, т. е. мы пренебрегаем большим числом отказов в силу малости соответствующих вероятностей. Появление x_k отказов на интервале времени Δ эквивалентно ситуации, когда мы на отрезок Δ бросили независимо x_k точек, каждая из которых распределена равномерно на этом интервале (рис. 9.4). Иначе говоря, условное распределение x_k точек пуассоновского потока на интервале Δ имеет многомерное равномерное распределение.

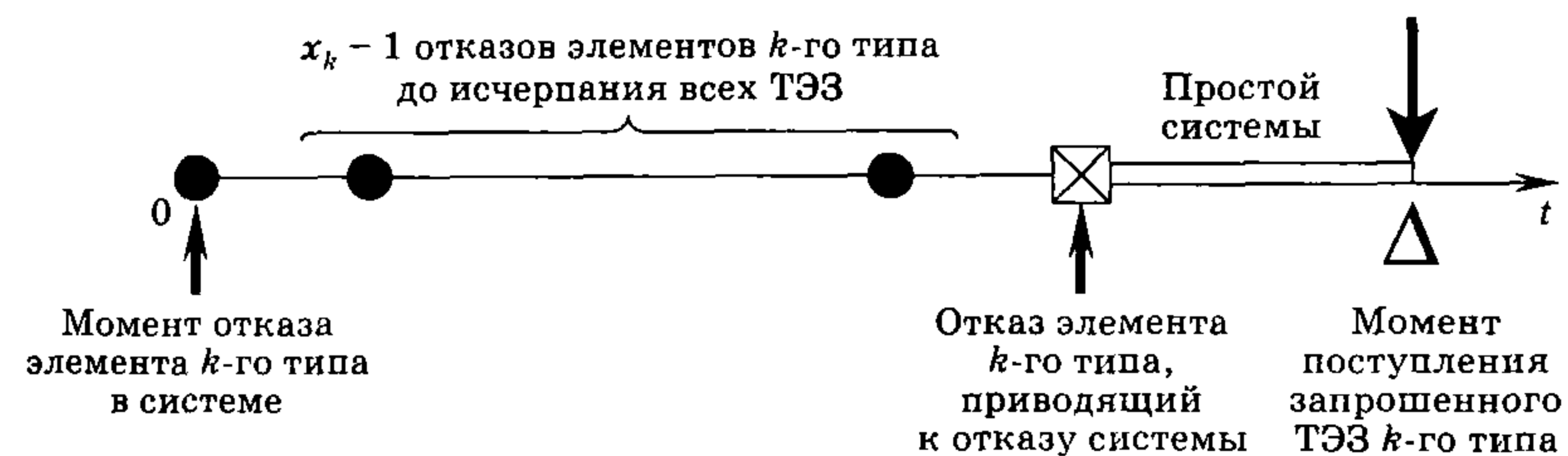


Рис. 9.4

Можно показать, что интервал Δ разбивается на $x_k + 1$ одинаково распределенных (хотя и зависимых) случайных интервалов, математическое ожидание каждого из которых составляет $\tau_k = \Delta / (x_k + 1)$.

Отказы элементов k -го типа в обслуживаемом объекте происходят с интенсивностью Λ_k , т. е. частота появления интервала простоя объекта из-за отказа элементов именно k -го типа равняется $\Lambda_k / \Lambda_{\text{сист}}$. Среднее время простоя объекта из-за нехватки ТЭЗ разных типов находится как средневзвешенное и равно

$$\tau_{\text{инд}} = \frac{1}{\Lambda_{\text{сист}}} \sum_{k=1}^n \tau_k \Lambda_k. \quad (9.7)$$

Коэффициент готовности индивидуального склада находится по стандартной формуле

$$K_{\text{инд}} = \frac{1}{1 + \Lambda_{\text{инд}} \tau_{\text{инд}}} = \frac{1}{1 + \sum_{k=1}^n \tau_k \Lambda_k} \approx 1 - \sum_{k=1}^n \tau_k \Lambda_k. \quad (9.8)$$

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО ЧИСЛА ТИПОВЫХ ЭЛЕМЕНТОВ ЗАМЕНЫ ЭЛЕМЕНТОВ НА ИНДИВИДУАЛЬНОМ СКЛАДЕ. Имея выражение для КГ и зная требование к индивидуальному складу по этому показателю, можно воспользоваться результатами гл. 7 для нахождения оптимального распределения запасных элементов, которое обеспечило бы минимально возможную стоимость склада.

§ 9.3. АНАЛИЗ РЕГИОНАЛЬНОГО СКЛАДА

ОБЩЕЕ ОПИСАНИЕ. Пусть региональный склад обслуживает M индивидуальных складов. Сам этот склад пополняется регулярно, через промежутки времени θ . Число хранимых на этом складе запасных ТЭЗ должно быть таковым, чтобы с заданной вероятностью ни один из индивидуальных складов не остался с неудовлетворенной заявкой. Поток заявок на требуемые ТЭЗ поступает в соответствии с происходящими отказами во всех обслуживаемых объектах, расположенных в зоне действия данного регионального склада.

Региональный склад содержит y_1, y_2, \dots, y_N ТЭЗ различных типов. Запас ТЭЗ каждого типа на региональном складе меньше, чем во всех индивидуальных складах региона, т. е. для каждого типа элементов выполняется условие

$$y_k \leq \sum_{j=1}^M x_{jk}, \quad (9.9)$$

где x_{jk} — число ТЭЗ k -го типа в j -м индивидуальном складе данного региона.

Задача заключается в том, чтобы найти такие оптимальные числа запасных ТЭЗ каждого типа y_1, y_2, \dots, y_N , которые бы обеспечили заданные требования на надежность регионального склада при минимальных затратах на запасные элементы. Для этого нам нужно найти формулу основной характеристики индивидуального склада — коэффициента готовности.

ПРАВИЛО ПОПОЛНЕНИЯ РЕГИОНАЛЬНОГО СКЛАДА. Региональные склады обычно пополняются периодически с некоторым фиксированным интервалом θ . Конечно, это не означает, что в случае преждевременного опустошения регионального склада невозможна экстренная доставка необходимых запчастей с центрального склада по запросу. Заметим, что запрос на экстренную поставку необходимых запасных ТЭЗ (или же осуществление преждевременной поставки в полном объеме) может осуществ-

ляться не только по исчерпанию соответствующего запаса полностью, но и в случае достижения некоторого заранее определенного порогового значения. Такой способ пополнения региональных складов минимизирует стоимость доставки необходимых запасных ТЭЗ. В противном случае схема сводилась бы фактически к работе центрального склада напрямую с индивидуальными, что привело бы к недопустимо большим издержкам.

Временная диаграмма, поясняющая возникновение простоя регионального склада при регулярном пополнении, приведена на рис. 9.5. Допустим, что заявка на экстренное пополнение посылается региональным складом на центральный склад, когда число наличных запасных ТЭЗ достигает некоторого порога. В примере, представленном на рисунке, номинальное число запасных ТЭЗ некоторого типа на региональном складе равно для конкретности пяти, а пороговый уровень равен единице, т. е. заявка высылается в момент, когда на складе остается только один элемент данного типа.

На первых двух циклах, а также на четвертом и шестом циклах исчерпания запаса ТЭЗ не происходит. На третьем при отсылке последнего элемента данного типа в центральный склад посылается заявка на экстренное пополнение. За время ожидания, равное Δ^* , нового отказа не происходит и региональный склад после пополнения до первоначального уровня продолжает благополучно функционировать. С момента пополнения начинается отсчет нового цикла регулярной длины θ .

На пятом цикле в интервале Δ^* происходит еще один отказ элемента того типа, запас которого уже исчерпался. Начинается простой регионального склада, τ^* , который заканчивается с прибытием пополнения с центрального склада.

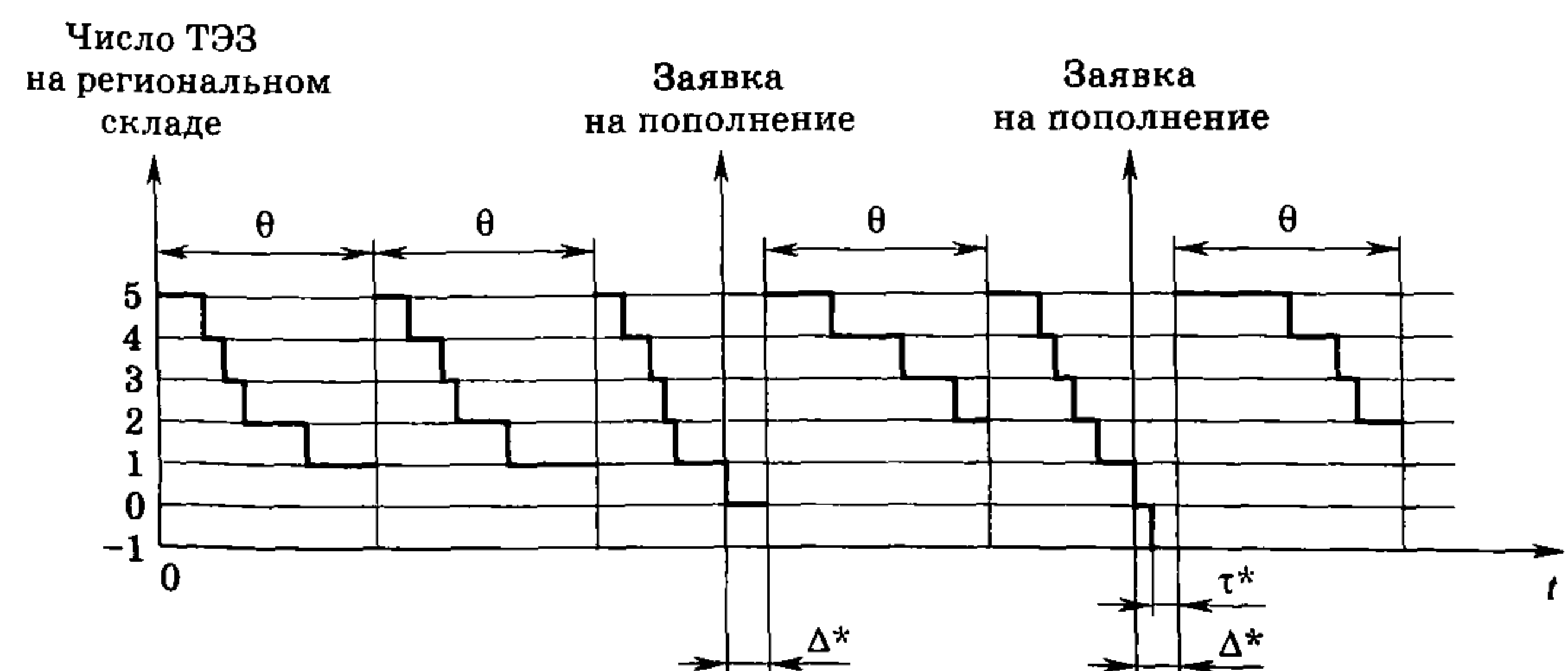


Рис. 9.5

ТРЕБОВАНИЯ К РЕГИОНАЛЬНОМУ СКЛАДУ. Региональный склад запасных ТЭЗ необходим для быстрой поставки затребованных элементов на индивидуальные склады. Чтобы определить оптимальное число запасных элементов каждого типа, необходимо сформулировать требования к надежности регионального склада. В этом случае, как и в предыдущем, можно базироваться на имеющемся практическом опыте проектирования и эксплуатации аналогичных складских систем. Однако, формулируя требования по надежности к региональному складу, можно опираться и на аналогичные требования к индивидуальному складу.

Пусть известно, что требуемое значение коэффициента готовности индивидуального склада равно $A_{инд}$. Опять возникает вопрос: какую степень ухудшения коэффициента готовности индивидуального склада мы готовы допустить за счет неидеальности обслуживания заявок региональным центром? Например, пусть коэффициент простоя индивидуального склада $U_{инд} = 0.0005$. Допустимо ли, что эта величина возрастет на сколько-то процентов?

Введем опять коэффициент β , показывающий, какое снижение надежности индивидуального склада допустимо за счет дефицита необходимых ТЭЗ на региональном складе. Для коэффициента простоя регионального склада $U_{рег}$ можно, очевидно, записать:

$$U_{рег} = \beta_{рег} U_{инд}. \quad (9.10)$$

Величина β в этом случае выбирается в диапазоне от 10% до 20% в зависимости от того, сколько индивидуальных складов обслуживается данным региональным складом. Поясним, почему эта цифра выше, чем коэффициент β для индивидуального склада, с помощью следующего примера. Пусть некоторый j -й индивидуальный склад исчерпал свой запас элементов k -го типа и в то же время послал очередной запрос на региональный склад. Однако на региональном складе также нет запасных элементов этого типа. В этом случае новый отказ элемента k -го типа в регионе приводит к отказу одного из региональных складов только при условии, что это будет отказ именно той же самой обслуживаемой системы, у которой индивидуальный склад исчерпал все свои запасные элементы данного типа. Вероятность этого неблагоприятного события равна

$$s_k = \frac{\lambda_{jk}}{\lambda_k}. \quad (9.11)$$

Это означает, что вероятность того, что отказ критического элемента приведет к отказу хотя бы одного индивидуального склада, фактически равна βs_k . Если в регионе K обслуживаемых систем, то можно считать, что $s_k \approx 1/K$, полагая, что все обслуживаемые системы в регионе приблизительно одинаковы по своему составу.

ИНТЕНСИВНОСТЬ ОТКАЗОВ РЕГИОНАЛЬНОГО СКЛАДА. Если на интервале θ на региональный склад со стороны индивидуальных складов данного региона поступит более чем y_k заявок на ТЭЗ k -го типа, то наступает отказ регионального склада. Пусть в j -м обслуживаемом объекте содержится m_{jk} элементов k -го типа, тогда общее число элементов этого типа, работающих в данном регионе, равно

$$m_k = \sum_{j=1}^M m_{jk}. \quad (9.12)$$

Вероятность \tilde{r}_k того, что отказ регионального склада не наступит из-за нехватки ТЭЗ k -го типа при наличии на нем y_k таких элементов, равна

$$\tilde{r}(y_k) = \sum_{j=0}^{y_k-1} \frac{(\lambda_k m_k \theta)^j}{j!} \cdot \exp(-\lambda_k m_k \theta), \quad (9.13)$$

где λ_k — интенсивность отказа элемента k -го типа, а θ — период планового пополнения регионального склада.

Для высоконадежных систем при $\tilde{q}_k(y_k) = 1 - \tilde{r}_k$ имеем

$$\tilde{q}_k(y_k) = \sum_{j=y_k}^{\infty} \frac{(\lambda_k m_k \theta)^j}{j!} \cdot \exp(-\lambda_k m_k \theta) \approx \frac{(\lambda_k m_k \theta)^{y_k}}{y_k!}. \quad (9.14)$$

Таким образом, \tilde{q}_k есть вероятность того, что отказ элементов k -го типа приведет к отказу регионального склада. Поток отказов элементов k -го типа по региону в целом является пуассоновским с параметром $\lambda_k m_k$, а поэтому поток отказов регионального склада из-за этих отказов равен

$$\tilde{\Lambda}_k = \lambda_k m_k \tilde{q}_k(y_k) \approx \lambda_k m_k \frac{(\lambda_k m_k \theta)^{y_k}}{y_k!}. \quad (9.15)$$

Результирующий поток отказов регионального склада составляет

$$\Lambda_{рег} = \sum_{k=1}^N \tilde{\Lambda}_k. \quad (9.16)$$

СРЕДНЕЕ ВРЕМЯ ПРОСТОЯ РЕГИОНАЛЬНОГО СКЛАДА. Рассмотрим ситуацию, когда региональный склад посылает заявку на экстренное пополнение по k -му типу элементов, как только уровень запасов ТЭЗ этого типа достигает порогового значения π_k . Задержка в поступлении экстренного пополнения равна Δ^* .

Допустим, что имевшийся на момент пополнения запас π_k исчерпался до прибытия пополнения. Это может произойти, если

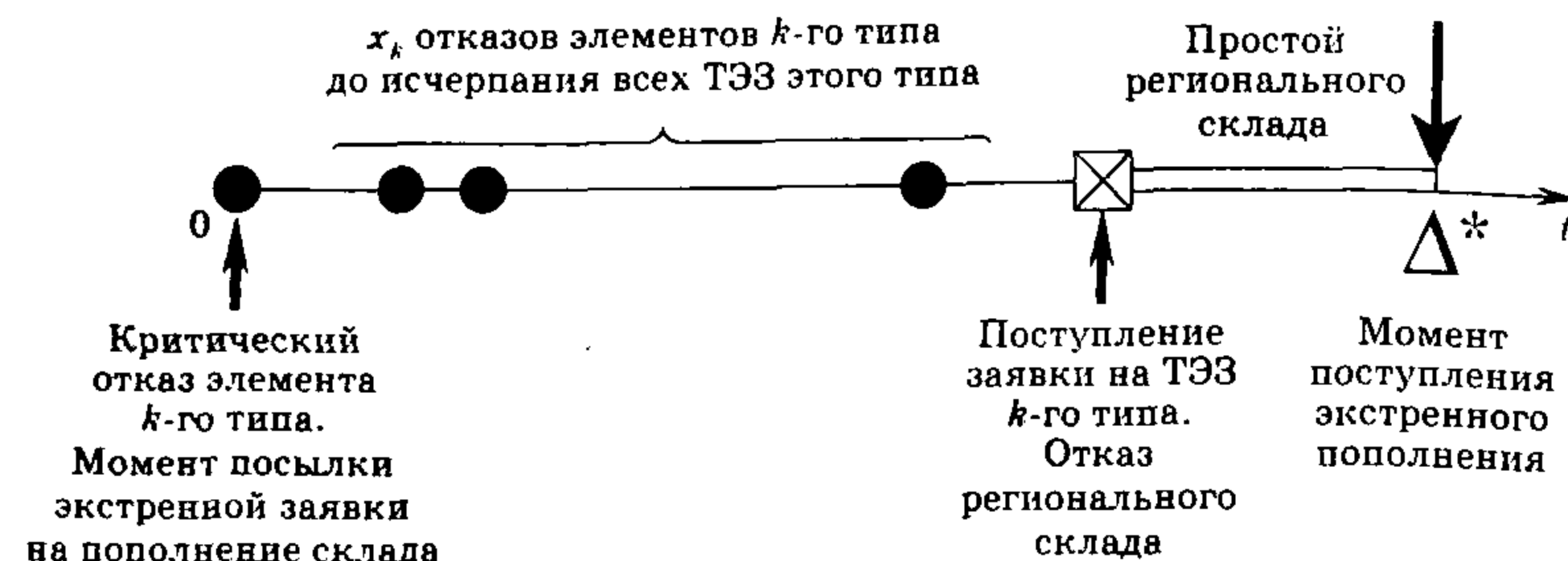


Рис. 9.6

число отказов, возникших в интервале Δ^* , будет равно или больше π_k , т. е. превысит число запасных ТЭЗ данного типа, оставшихся на региональном складе. Рассматривая высоконадежную систему, мы можем сделать предположение, что в интервале Δ^* возникло π_k отказов, т. е. пренебрежем возможностью появления большего числа отказов. Это означает, что на отрезок Δ^* независимо брошено π_k точек, каждая с равномерным распределением внутри этого интервала (рис. 9.6).

Следовательно, интервал Δ^* разбивается на $(\pi_k + 1)$ одинаково распределенных случайных интервалов, математическое ожидание каждого из которых равно

$$\tau_{\text{рег}} = \frac{\Delta^*}{\pi_k + 1}. \quad (9.17)$$

Если положить для всех типов ТЭЗ один и тот же пороговый уровень π , то среднее время простоя регионального склада составляет

$$\tau_{\text{рег}} = \frac{\Delta^*}{\pi + 1}. \quad (9.18)$$

КОЭФФИЦИЕНТ ГОТОВНОСТИ РЕГИОНАЛЬНОГО СКЛАДА находится по стандартной формуле

$$K_{\text{рег}} = \frac{1}{1 + \Lambda_{\text{рег}} \tau_{\text{рег}}} \approx 1 - \Lambda_{\text{рег}} \tau_{\text{рег}}. \quad (9.19)$$

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО ЧИСЛА ЭЛЕМЕНТОВ НА РЕГИОНАЛЬНОМ СКЛАДЕ. Имея выражение для КГ и зная требование к индивидуальному складу по этому показателю, используем результаты гл. 7 для нахождения оптимального распределения запасных элементов, которое обеспечило бы минимально возможную стоимость запасных элементов, хранящихся на складе.

§ 9.4. ЦЕНТРАЛЬНЫЙ СКЛАД

ОБЩЕЕ ОПИСАНИЕ. Объем и номенклатура центрального склада зависит от общего числа обслуживаемых объектов, структуры региональных складов и производительности ремонтной базы. Кроме того, часть ТЭЗ после отказа становятся неремонтопригодными и их приходится заменять на новые. Эти новые ТЭЗ доставляются на центральный склад соответствующими поставщиками. Опустим возможность появления «летальных» отказов сменных элементов, т. е. появления таких отказов, которые делают невозможным ремонт ТЭЗ.

В этом случае центральный склад является своеобразным буферным аккумулятором отремонтированных ТЭЗ, поступающих с ремонтной базы. Этот буфер играет роль *сглаживающего фильтра*: часть ТЭЗ в момент необходимого пополнения региональных складов находится в ремонте. Если бы ремонт осуществлялся практически мгновенно, то отремонтированных ТЭЗ вполне хватало бы на восполнение запасов региональных складов. Дополнительный запас ТЭЗ на центральном складе служит для того, чтобы компенсировать задержки на ремонтной базе. Возникает проблема: стоит ли увеличить запас лишних ТЭЗ на центральном складе или же увеличить пропускную способность ремонтной базы.

Если пренебречь экстренными пополнениями региональных складов, которые в высоконадежной складской системе должны происходить достаточно редко, то запас на центральном складе должен быть таковым, чтобы с заданной гарантией покрыть число ТЭЗ, находящихся на ремонтной базе непосредственно в стадии ремонта или в процессе ожидания.

Центральный склад содержит z_1, z_2, \dots, z_n ТЭЗ различных типов. Пусть система включает N регионов со своими складами. Задача заключается в том, чтобы найти такие оптимальные числа запасных ТЭЗ каждого типа z_1, z_2, \dots, z_N , которые бы обеспечили заданные требования на надежность центрального склада при минимальных затратах на запасные элементы.

ПРАВИЛО ПОПОЛНЕНИЯ ЦЕНТРАЛЬНОГО СКЛАДА. Центральный склад накапливает отремонтированные ТЭЗ, поступающие с ремонтной базы. Если за текущий период отказало много элементов в объектах системы, то с ремонтной базы на центральный склад поступит много отремонтированных ТЭЗ. Однако и заявок на пополнение региональных складов будет больше в момент следующей поставки. Если бы ремонт осуществлялся практически мгновенно, то центральный склад не должен был бы иметь никаких

дополнительных ТЭЗ. Однако в силу конечной пропускной способности ремонтной базы часть ТЭЗ будет постоянно в том или ином количестве находиться на ремонтной базе.

Если отказом центрального склада можно назвать состояние, когда он не может удовлетворить всех заявок на поставку необходимых ТЭЗ во все региональные склады, то должен быть дополнительный страховой запас, который компенсировал бы задержку ремонта. Нужно заметить, что подобный критерий отказа является весьма жестким: недопоставка ТЭЗ во время периодического пополнения региональных складов не так уж существенна, если имеется возможность дослать недостающие ТЭЗ несколько позже.

ТРЕБОВАНИЯ К ЦЕНТРАЛЬНОМУ СКЛАДУ И РЕМОНТНОЙ БАЗЕ. Центральный склад запасных ТЭЗ необходим для пополнения региональных складов. Сам центральный склад жестко привязан к ремонтной базе: по существу, ремонтная база является источником поступлений отремонтированных ТЭЗ. В связи с этим требования по надежности следует предъявлять к паре: центральный склад — ремонтная база.

Как и в предыдущих случаях, допустимо основываться на имеющемся практическом опыте. Однако требования по надежности к центральному складу можно задавать, основываясь на существующих требованиях к региональному складу.

Если известно значение КГ регионального склада $A_{\text{рег}}$, то задают допустимый коэффициент $\beta_{\text{центр}}$, $\beta_{\text{центр}} < 1$. Тогда коэффициент простоя регионального склада $U_{\text{центр}}$ имеет вид

$$U_{\text{центр}} = \beta_{\text{центр}} U_{\text{рег}}. \quad (9.20)$$

Таким образом, получаем выражение для КГ центрального склада:

$$A_{\text{центр}} = 1 - \beta_{\text{центр}}(1 - A_{\text{рег}}). \quad (9.21)$$

Величина $\beta_{\text{центр}}$ может быть выбрана порядка 10% благодаря меньшей критичности КГ центрального склада: в случае текущей нехватки ТЭЗ на запросы с региональных складов недостающие ТЭЗ могут быть досланы позже.

РАСПРЕДЕЛЕНИЕ ЧИСЛА ТЭЗ, НАХОДЯЩИХСЯ ОДНОВРЕМЕННО НА РЕМОНТНОЙ БАЗЕ. Ремонтная база представляет собой многоканальную систему массового обслуживания. Производительность (скорость ремонта) зависит в основном от двух факторов: числа ремонтных органов, работающих одновременно, и интенсивности

проведения ремонтных работ каждым из них. Нас интересует распределение числа отказавших ТЭЗ, находящихся одновременно в ремонтной базе. Это связано с тем, что именно число отказавших элементов в ремонтных органах определяет, сколько избыточных запасных ТЭЗ должно иметься на центральном складе.

Ремонтная база и центральный склад тесно связаны между собой: чем меньше ремонтных органов на ремонтной базе, т. е. чем дольше обслуживание, тем больше нужен запас избыточных ТЭЗ на центральном складе. Иначе говоря, проблему производительности ремонтной базы нужно решать совместно с проблемой необходимого объема ТЭЗ на центральном складе.

В предположении экспоненциальности всех распределений ремонт ТЭЗ может быть описан при помощи марковского процесса *гибели и размножения* (см. Приложение 4).

На ремонтную базу поступает результирующий поток $\tilde{\Lambda}$ отказов элементов всех типов. С помощью модели, приведенной в Приложении 5, можно найти стационарное распределение числа элементов k , находящихся на ремонтной базе в некий произвольный момент времени. Обозначим эту вероятность через p_k . Найдем номер состояния v , удовлетворяющий условию

$$\sum_{k=v+1}^{\infty} p_k \leq U_{\text{центр}}. \quad (9.22)$$

Если бы все элементы были идентичны, то v определяло бы число элементов, которые необходимо иметь на центральном складе. Однако суммарный поток отказов состоит из k подпотоков различных типов элементов с интенсивностями $\tilde{\Lambda}_k$. В среднем среди отказавших элементов, находящихся на ремонтной базе, элементы k -го типа составляют долю $\tilde{\Lambda}_k/\tilde{\Lambda}$. На основании этого можно сделать грубую оценку числа необходимых ТЭЗ каждого типа на центральном складе:

$$v_k = v \frac{\tilde{\Lambda}_k}{\tilde{\Lambda}}. \quad (9.23)$$

Такая оценка является приемлемой, поскольку средние значения числа запасных элементов на центральном складе велики. Более точная оценка числа ТЭЗ каждого типа v_k может быть проведена с использованием полиномиального распределения, но из-за громоздкости здесь не приводится. Решение (9.23) является практически приемлемым.

§ 9.5. ПРОГРАММА ВЫБОРА ОПТИМАЛЬНОГО ЧИСЛА ЗАПАСНЫХ ЭЛЕМЕНТОВ ДЛЯ СПУТНИКОВОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ* «GLOBALSTAR»

Несколько десятков наземных базовых станций спутниковой телекоммуникационной сети «GlobalStar» разбросаны по всему земному шару. Каждая наземная станция имеет свою собственную конфигурацию (число принимающих и передающих каналов, буферная память, коммутационная часть и т. п.), хотя состоит из идентичных типовых элементов. Для оперативной замены отказавших элементов на соответствующие типовые элементы замены (ТЭЗ) каждая наземная станция снабжена своим индивидуальным складом. Сами эти индивидуальные склады снабжаются с региональных складов после каждой очередной замены отказавшего элемента.

Для расчета оптимального числа типовых элементов замены на индивидуальных и региональных складах была разработана специальная компьютерная программа**, краткое описание пользовательского интерфейса которой приводится ниже. В расчетном блоке программы использовался метод квот, описанный в гл. 8.

Основное программное окно представлено на рис. 9.7. В нем отражены все используемые в программе операторские команды. Наиболее часто используемые команды верхнего уровня выведены в виде клавиш на верхнем уровне окна. Здесь же отражена иерархическая структура складской системы — центральный, региональные и индивидуальные склады.

В процессе определения оптимального числа ТЭЗ на складе выполняются следующие шаги.

■ Формируется список используемых ТЭЗ с указанием стоимости и интенсивности отказов для каждого из них (рис. 9.8). Здесь в столбцах указаны:

- ✓ идентификационный номер элемента;
- ✓ название элемента;
- ✓ среднее время безотказной работы;
- ✓ стоимость;
- ✓ дополнительная информация, например источник исходных данных.

* Все логотипы и торговые марки, упоминаемые в тексте, принадлежат их владельцам.

** Компьютерная программа «Optimal Spare Allocation» разработана С. В. Антоновым и автором.

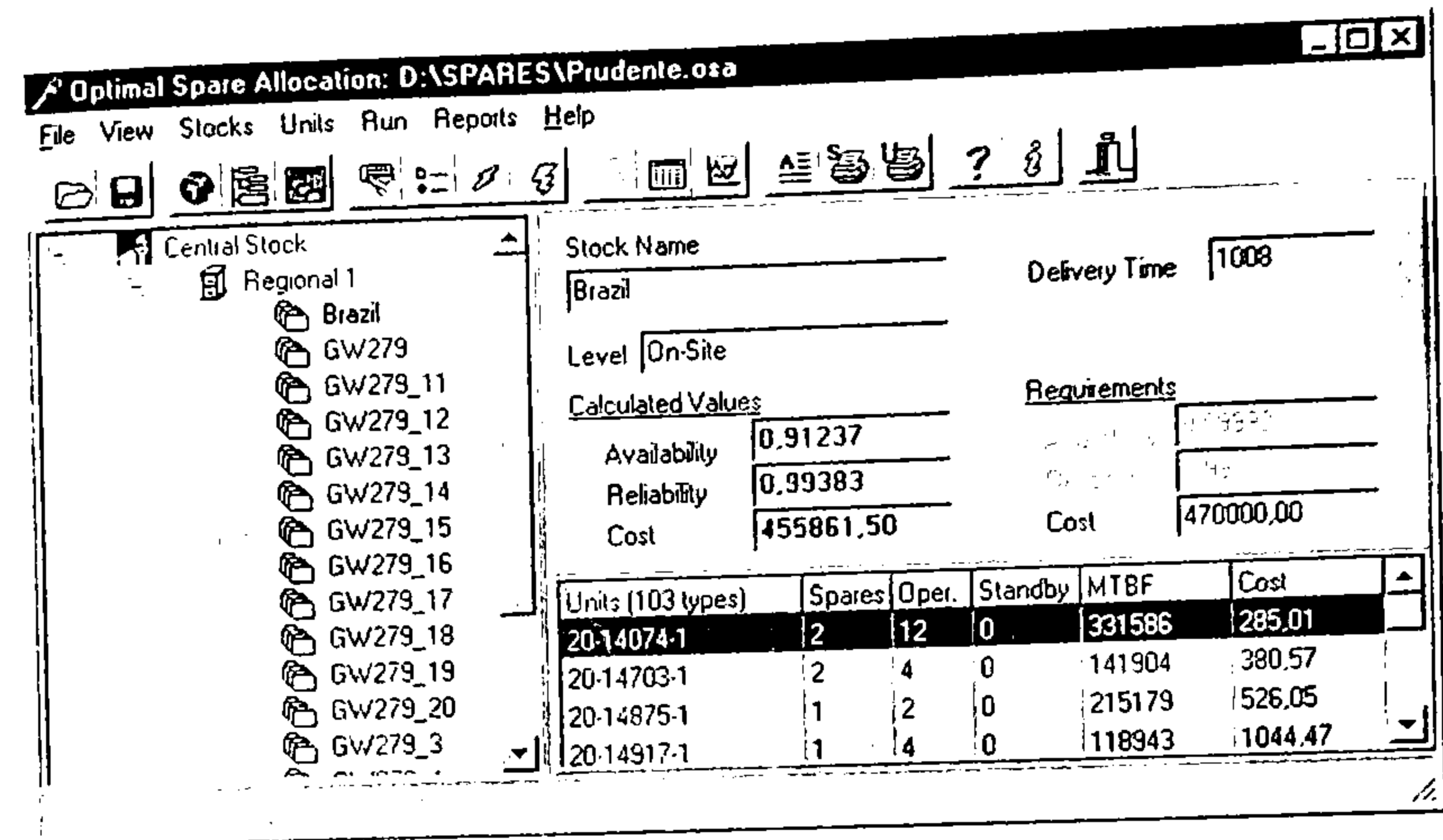


Рис. 9.7

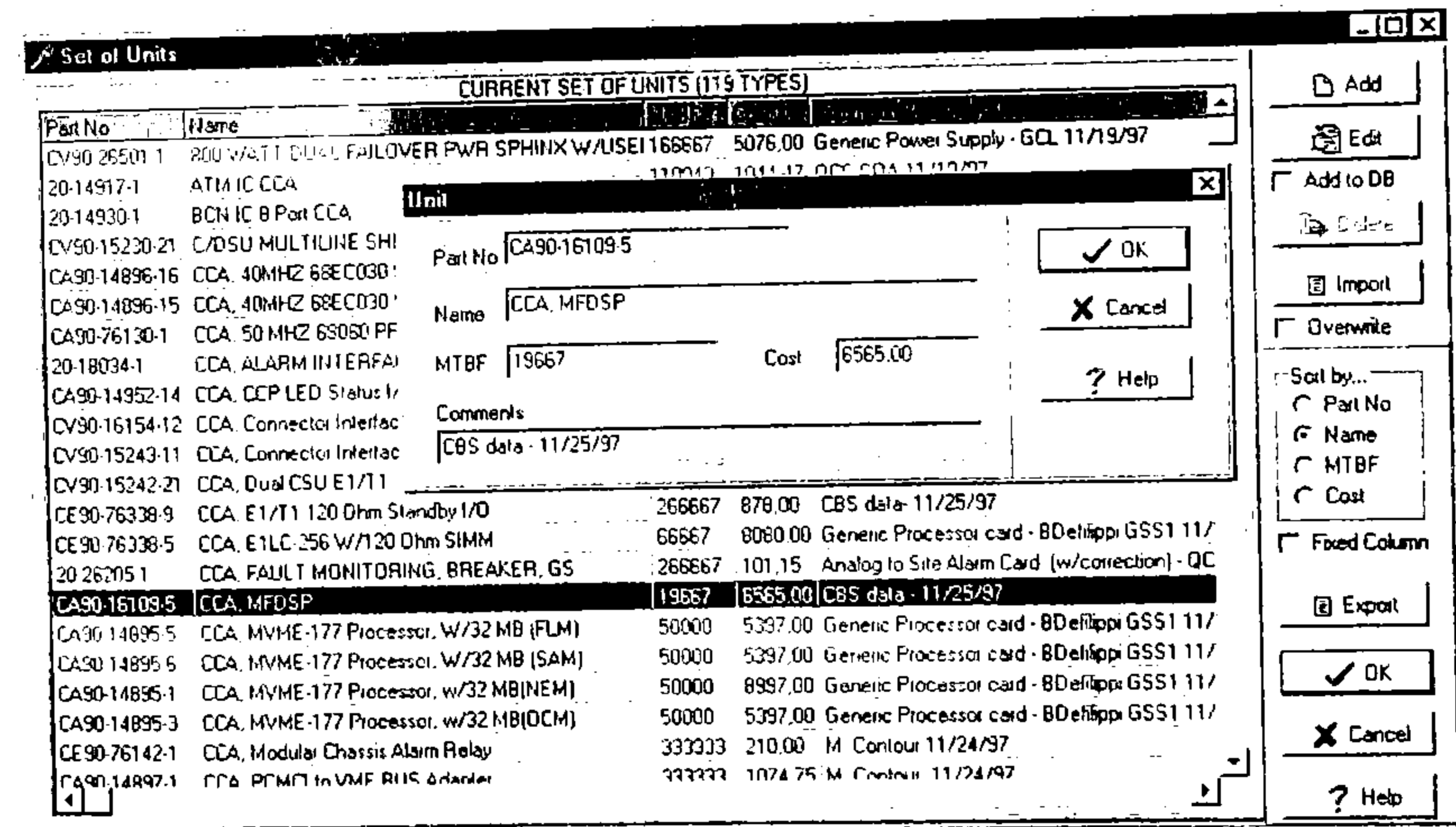


Рис. 9.8

Внутри большого окна в малом окне показываются атрибуты выделенного элемента в большом окне. Это окно может быть использовано для корректировки исходных данных.

■ Формируется иерархическая структура складов. Структура может быть неоднородной: некоторые индивидуальные склады могут

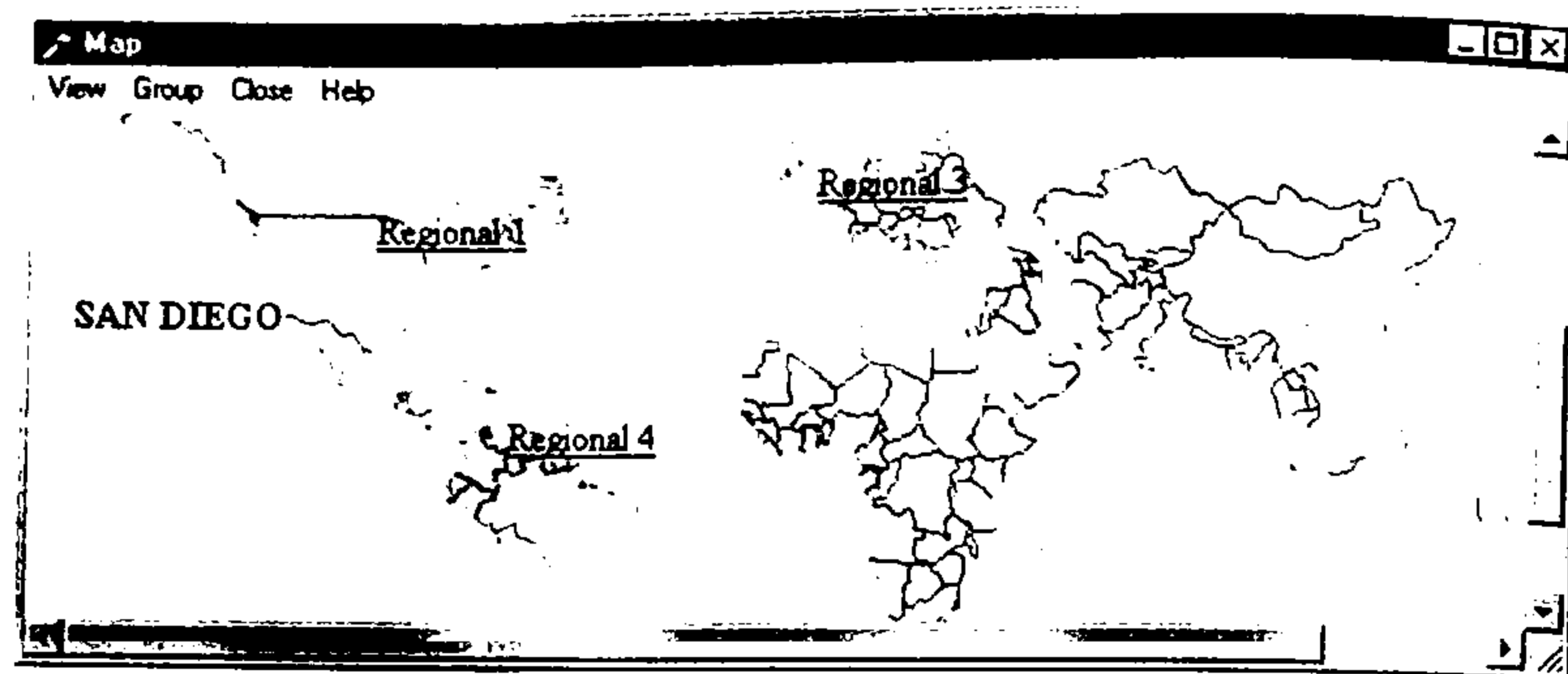


Рис. 9.9

снабжаться напрямую с центрального склада. Формирование иерархической структуры позволяет агрегировать на региональных складах информацию о числе обслуживаемых элементов каждого типа индивидуальными складами. Сформированная иерархическая структура складов видна на рис. 9.7. Эта же структура отображается на карте: каждый склад любого уровня может быть вызван непосредственно с карты. На рис. 9.9 представлен уровень региональных складов.

- Формируются списки типовых элементов, используемых в каждой наземной базовой станции с указанием их числа. Окна, соответствующие этому шагу, представлены на рис. 9.10. Пер-

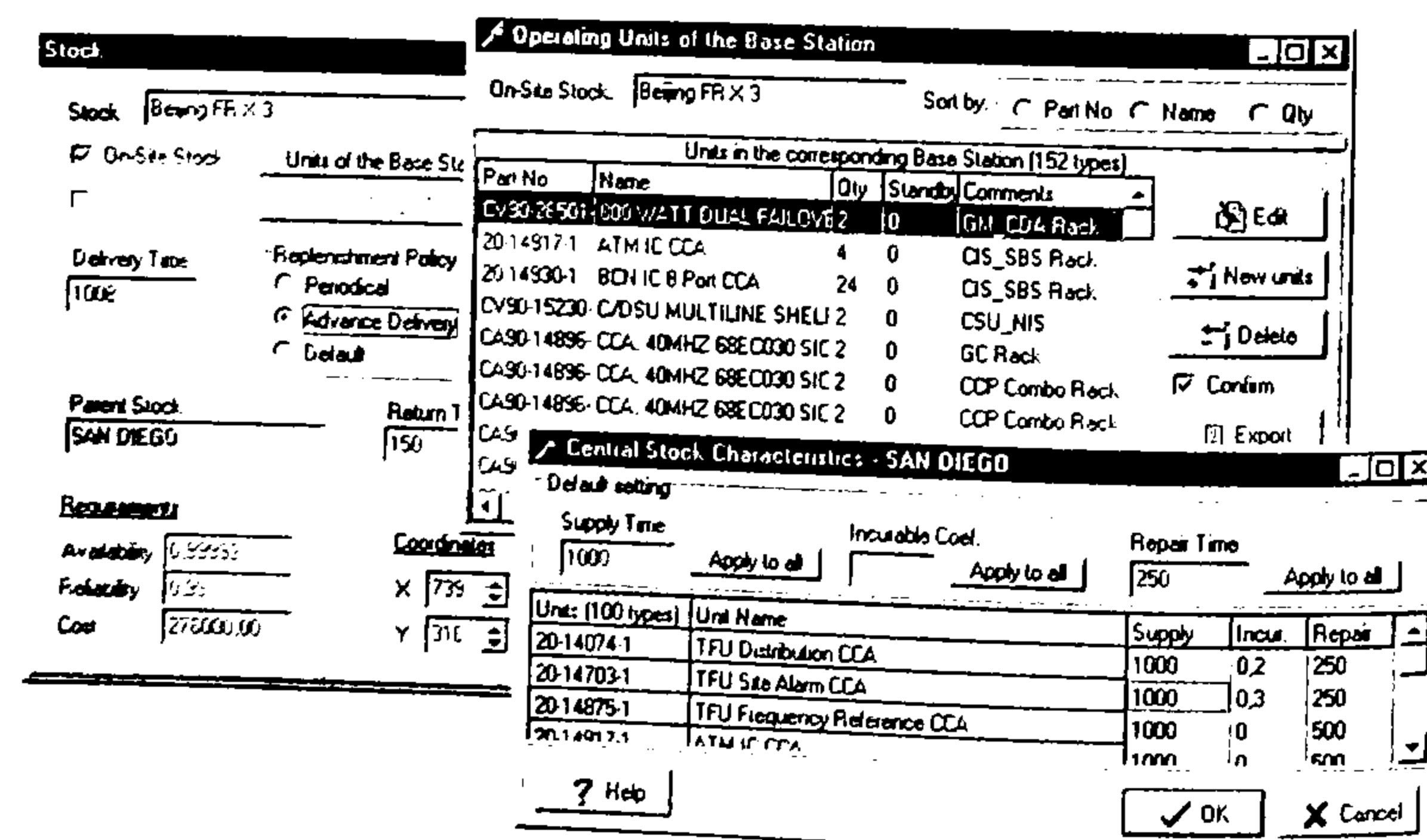


Рис. 9.10

вое окно (верхний слой) показывает характеристики центрального склада (по колонкам):

- ✓ идентификационный номер элемента;
- ✓ название элемента;
- ✓ время доставки на региональный склад;
- ✓ коэффициент, характеризующий ремонтпригодность отказавшего ТЭЗ данного типа;
- ✓ среднее время ремонта отказавшего ТЭЗ.

Окно на следующем уровне содержит следующую информацию:

- ✓ идентификационный номер типового элемента;
- ✓ название элемента;
- ✓ число элементов данного типа в наземной станции, обслуживаемой данным индивидуальным складом;
- ✓ число встроенных резервных элементов — *горячий резерв*;
- ✓ название стойки аппаратуры, где размещены данные элементы.

Наконец, окно на нижнем уровне содержит следующие данные:

- ✓ время доставки ТЭЗ на региональный склад;
- ✓ тип доставки (по запросу, регулярно);
- ✓ требования к складу (коэффициент готовности, вероятность бесперебойного снабжения, стоимость);
- ✓ координаты склада для отображения на географической карте.

- Выбираются параметры для расчета. Соответствующее этому шагу окно представлено на рис. 9.11. В левой колонке представлены:

- ✓ возможность выбора целевой функции (вероятность истощения запаса ТЭЗ хотя бы одного типа);
- ✓ используемый метод расчета (метод квот, метод наискорейшего спуска, ручная корректировка);
- ✓ тип задачи (минимизация стоимости или максимизация надежности склада);
- ✓ вспомогательные характеристики (выдача последовательности решений, дополнительная субоптимизация при исключении крайне дорогостоящих элементов, точность вычислений).

Правая колонка содержит:

- ✓ правило пополнения склада (периодически или по запросу в момент замены);
- ✓ требуемое начальное состояние склада (нет запаса, обязательное априорное наличие одного элемента каждого типа, заданное директивно);
- ✓ некоторые характеристики центрального склада (ремонтпригоден ли ТЭЗ или всегда требует замены на новый).

Глава 10

Территориально-распределенные системы технического обслуживания

Рассматривается построение оптимальных зон технического обслуживания территориально-распределенных систем передвижными ремонтными мастерскими (ПРМ). Изучается задача определения оптимального числа этих ПРМ в каждой зоне, приводится эвристическое ее решение и дается практический пример.

§ 10.1. ОПИСАНИЕ АНАЛИЗИРУЕМОЙ СИСТЕМЫ

По территории, разбитой на районы (например, административные), распределены некоторым известным образом технические объекты, в случае отказов которых посылаются заявки на обслуживание. Известна интенсивность поступления заявок от каждого района, определенная средним числом отказов в единицу времени.

Предполагается, что одна передвижная ремонтная мастерская (ПРМ) может обслуживать несколько районов, хотя при высокой плотности распределения объектов может потребоваться несколько ПРМ для одного района. Работа всех ПРМ в данной зоне обслуживания координируется их базой.

Задача заключается в построении зон обслуживания таких, чтобы общее число ПРМ на всей территории было бы минимальным при условии, что качество обслуживания удовлетворяет заданным требованиям. Качество обслуживания будем характеризовать двумя основными факторами:

- ✓ задержкой подтверждения со стороны базы ПРМ того, что заявка на обслуживание принята и соответствующая ПРМ послана на место; такая задержка может быть вызвана занято-

стью всех ПРМ обслуживанием других объектов в момент поступления заявок;

- ✓ временем обслуживания, включающим в себя время, которое ПРМ проводит в пути до отказавшего объекта и временем собственно ремонтных процедур на месте.

Проведем некоторые рассуждения. С увеличением радиуса зоны обслуживания коэффициент использования ПРМ возрастает, однако в то же время ухудшается качество обслуживания: в частности возрастает

- ✓ возможность ожидания начала обслуживания из-за очереди;
- ✓ время в пути до отдаленных объектов.

Для группы соседних районов с большим числом клиентов возможно создание одной зоны обслуживания с несколькими ПРМ.

Возникает задача определения оптимальных размеров зон обслуживания и вычисление оптимального количества оперирующих ПРМ в них при условии удовлетворения заданным требованиям по качеству.

§ 10.2. ЗОНА С ЕДИНСТВЕННОЙ ПЕРЕДВИЖНОЙ РЕМОНТНОЙ МАСТЕРСКОЙ

10.2.1. Максимальный размер зоны

В случае низкой интенсивности отказов на единицу площади размер зоны обслуживания, в первую очередь, определяется физической возможностью ПРМ достижения отказавшего объекта за допустимое время. Например, для восьмичасового рабочего дня при времени ремонта, равном двум часам, имеется только шесть часов на дорогу к объекту и обратно, даже если заявка на обслуживание принята с самого утра*.

Допустим, что средняя скорость передвижения ПРМ равна 40 км/ч. Это означает, что максимальный радиус зоны обслуживания составляет около 120 км (три часа езды в один конец). В этом случае ПРМ способна обслужить заявку от далеко распо-

* Фактически рабочий день ПРМ не бывает столь жестко регламентирован, однако это не учитывается нами ради упрощения изложения.

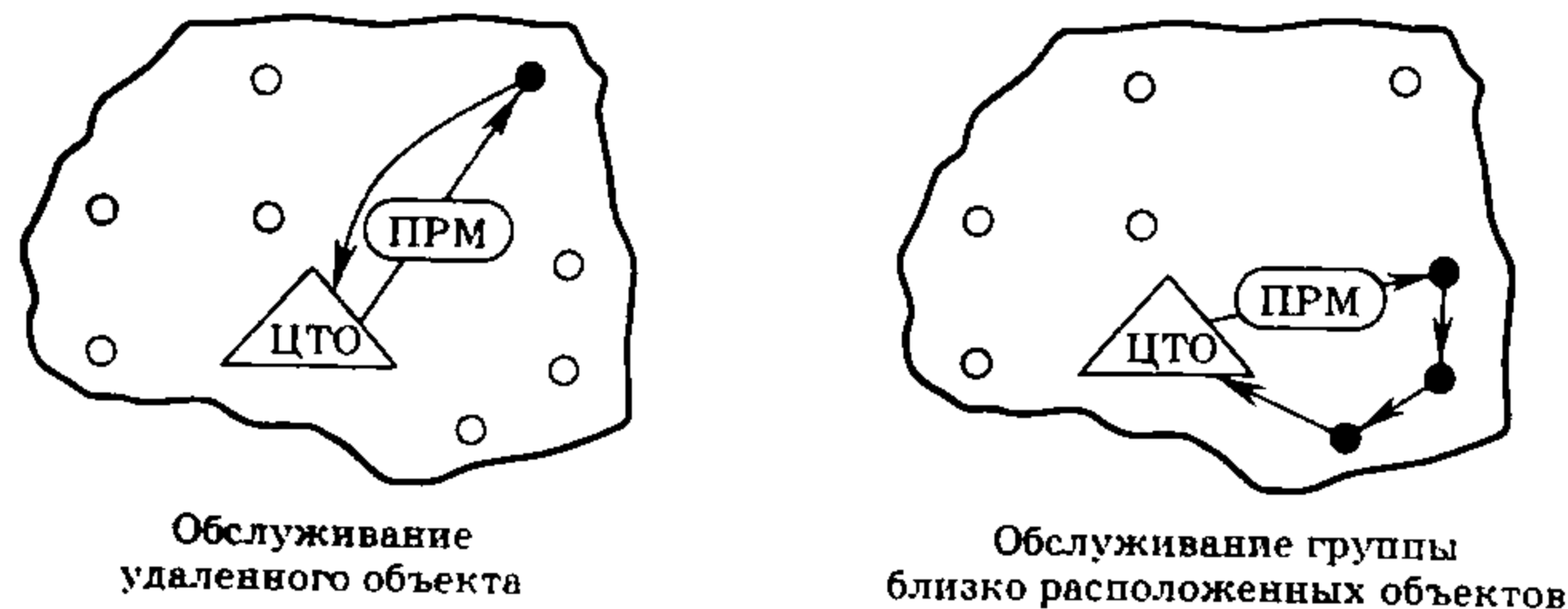


Рис. 10.1

ложенного объекта, если эта заявка поступила утром. Конечно, заявка, поступившая в середине или в конце дня, может быть обслужена только на следующий день. В то же время возможно обслуживание нескольких заявок, поступивших одновременно от близко расположенных объектов.

Говоря о низкой интенсивности запросов мы имеем в виду, что вероятность появления в зоне более одного запроса в сутки относительно мала. Обычно в таких случаях ПРМ бывает существенно недогружена и поэтому из экономических соображений должно быть предусмотрено ее использование для обслуживания фоновых задач.

Пример зоны обслуживания единственной ПРМ представлен на рис. 10.1. Здесь светлыми кружочками обозначены исправные, а темными — неисправные объекты, ЦТО — центр технического обслуживания.

10.2.2. Модель массового обслуживания для зоны с единственной передвижной ремонтной мастерской

Пусть Λ есть интенсивность запросов в зоне обслуживания, а μ — скорость (интенсивность) обслуживания, которую можно приблизительно определить как

$$\mu = \frac{1}{\text{среднее время доставки} + \text{среднее время ремонта}}. \quad (10.1)$$

ПРМ может быть описана системой массового обслуживания типа $M/M/1$.

* $M/M/1$ — стандартное в теории массового обслуживания обозначение системы с пуассоновским входящим потоком и экспоненциальным временем обслуживания при одном обслуживающем канале.

Введем безразмерный коэффициент нагрузки этой системы массового обслуживания в виде

$$\rho = \frac{\Lambda}{\mu}.$$

Среднее время ожидания в такой системе*

$$W = \frac{\Lambda}{\mu(\mu - \Lambda)}. \quad (10.2)$$

Заметим, что и частота поступления заявок (т. е. интенсивность отказов объектов) Λ , и скорость обслуживания μ зависят от радиуса зоны обслуживания r . Для зоны с радиусом r и равномерным распределением объектов по площади можно записать выражение для суммарной интенсивности заявок в зоне как

$$\Lambda(r) = \lambda \pi r^2, \quad (10.3)$$

где λ — интенсивность запросов в час на единицу площади, а πr^2 — площадь зоны в тех же единицах. Суммарная интенсивность заявок в зоне может быть задана и на основании реальных данных, полученных во время предварительного опыта обслуживания в данной зоне.

Средняя интенсивность обслуживания объекта в зоне равна

$$\mu(r) = \frac{1}{\tau + r^*/v}, \quad (10.4)$$

где τ — среднее время ремонта, r^* — среднее расстояние до объекта в зоне, а v — скорость перемещения ПРМ. Величина r^* зависит от размещения обслуживаемых объектов по отношению к базе ПРМ.

Окончательно формула для времени ожидания начала обслуживания запишется в виде

$$W = \frac{\lambda \pi r^2 (r^*/v + \tau)}{(r^*/v + \tau)^{-1} - \lambda \pi r^2}. \quad (10.5)$$

ПРИМЕР 10.1. Приведем на рис. 10.2 ту же топографию размещения объектов и базы ПРМ, что и на рис. 10.1. Представим карту в формализованном для решения задачи виде. Здесь через \mathcal{A} обозначена ПРМ, а через a, b, c, \dots, h — объекты обслуживания.

Заметим, что для транспортных задач так называемая «манхэттенская метрика»** является более подходящей, чем евк-

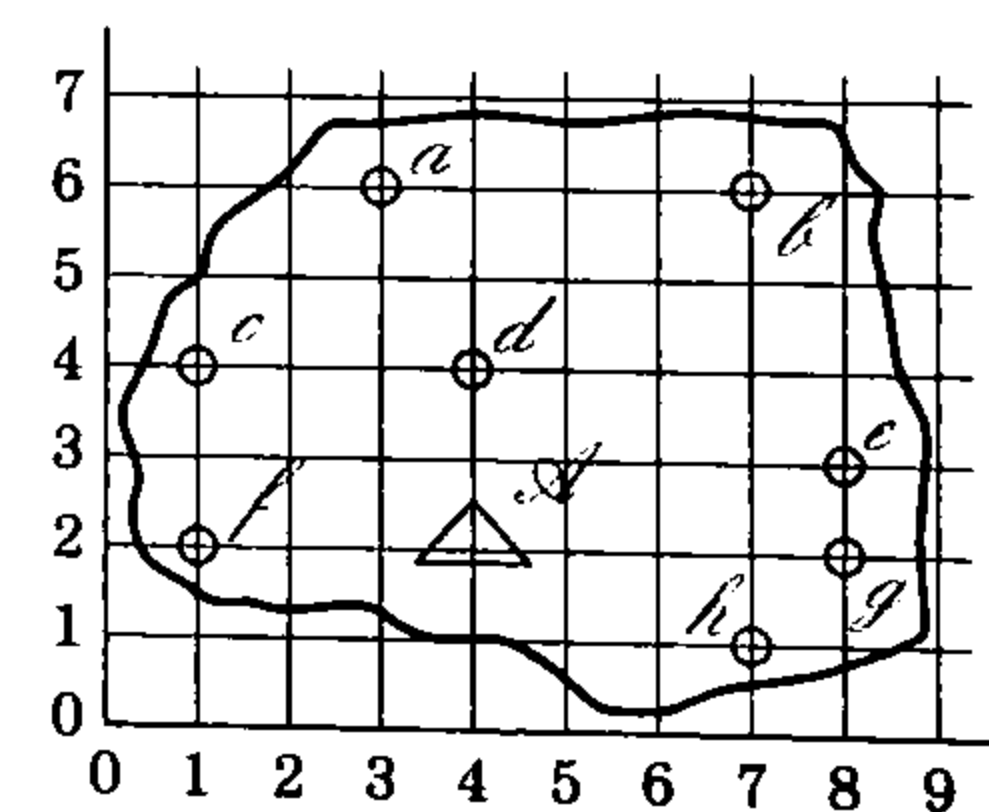


Рис. 10.2

* Вывод формулы (10.2) приведен в Приложении 5.

** Манхэттенской метрикой называется метрика в пространстве, в котором разрешено только движение по линиям «Север — Юг» и «Запад — Восток» аналогично движению на Манхэттене в Нью-Йорке.

лидова, поскольку дороги не являются прямыми линиями, соединяющими пункты. В соответствии с манхэттенской метрикой расстояние между пунктами с координатами (x_1, y_1) и (x_2, y_2) равняется $|x_1 - x_2| + |y_1 - y_2|$. Это несколько облегчает вычисление расстояний между базой ПРМ и обслуживаемыми объектами. В данном случае расстояния равны (в условных единицах масштаба рисунка): $Ab = 5$, $Ac = 7$, $Ad = 5$, $Ae = 2$, $Af = 5$, $Ag = 3$, $Ah = 4$, а среднее расстояние равно, таким образом, 4.375.

§ 10.3. ЗОНА ОБСЛУЖИВАНИЯ С НЕСКОЛЬКИМИ ПЕРЕДВИЖНЫМИ РЕМОНТНЫМИ МАСТЕРСКИМИ

10.3.1. Общее описание

Рассмотрим ту же зону обслуживания, что и ранее, предположив, что, например, интенсивность запросов возросла, потребовав увеличения числа ПРМ (рис. 10.3).

Физическое ограничение на размер зоны (радиус в 120 км) сохраняется. Однако предполагается, что число запросов в зоне может оказаться таким, что для технического обслуживания объектов будет недостаточно единственной ПРМ в данной зоне.

Моделью обслуживания объектов группой ПРМ в этой зоне может быть n -канальная система массового обслуживания типа $M/M/n^*$ с дисциплиной обслуживания FIFO**.

Среднее время обслуживания может быть записано как (см. Приложение 5, (П 5.23))

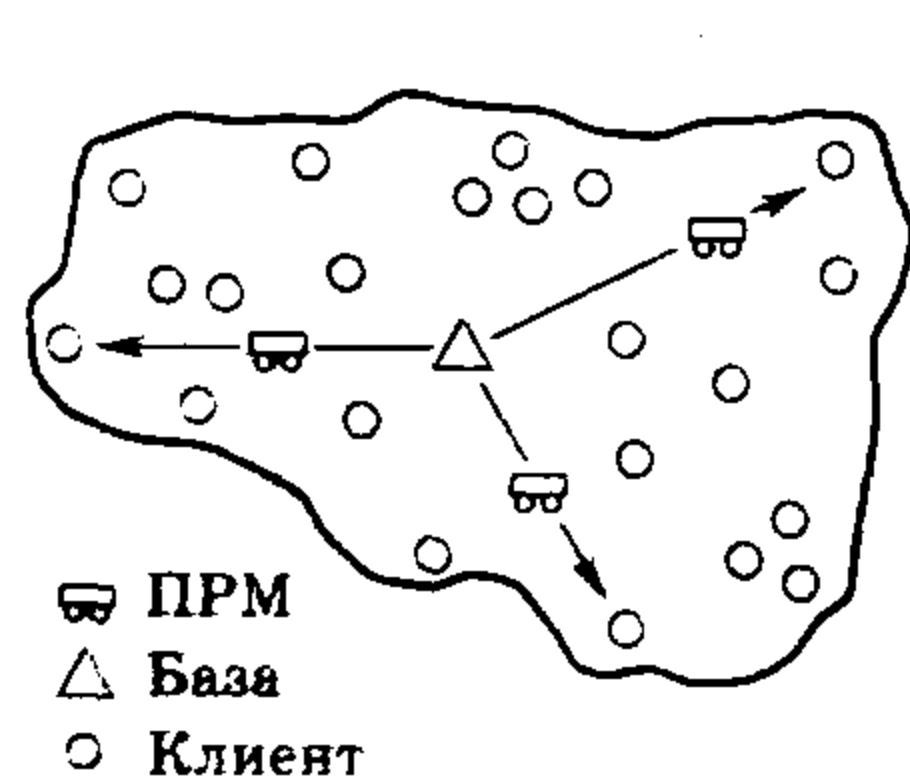


Рис. 10.3

$$W = \left[\frac{\rho^n \mu}{(n-1)!(n\mu - \lambda)^2} \right] p_0, \quad (10.6)$$

где

$$p_0 = \frac{1}{1 + \sum_{j=1}^k \frac{\rho^j}{j!} + \frac{\rho^{k+1}}{k!(k-\rho)}} \quad (10.7)$$

есть стационарная вероятность того, что n -канальная система обслуживания свободна (в данный момент не обслуживается ни одной заявки).

* $M/M/n$ — стандартное обозначение для n -канальной системы массового обслуживания с пуассоновским входным потоком и экспоненциально распределенным временем обслуживания.

** FIFO — дисциплина обслуживания в порядке поступления (аббревиатура «First In — First Out»).

10.3.2. Краткое описание метода

Предлагаемая процедура отыскания оптимального размера зон обслуживания определения необходимого числа ПРМ и размещения их базы в каждой из зон является многошаговой итеративной и существенно интерактивной процедурой. На каждом шаге процесса необходимо принимать решения в зависимости от уже полученных текущих результатов. Процедура эта в общих чертах следующая.

- (1) Построить изолированную оптимальную зону для одиночной ПРМ, используя ранее описанный метод анализа зон с одиночной ПРМ.
- (2) Построить соседнюю изолированную зону.
- (3) Проверить, можно ли агрегировать эти две зоны в одну с двумя ПРМ, приняв во внимание характеристики обслуживания: время ожидания и время в пути.
- (4) Построить следующую соседнюю зону. Добавление зон должно идти таким образом, чтобы в процессе *разрастания* общая территория, покрытая зонами обслуживания, имела бы по возможности компактную «круглую» форму.
- (5) Повторить процедуру с шага (3).
- (6) Продолжать процедуру до тех пор, пока вся территория не будет покрыта зонами обслуживания.

Заметим, что целевая функция является многоэкстремальной, т. е. описанное выше итеративное решение может существенно зависеть от начально выбранной *точки разрастания* зоны обслуживания. Вообще говоря, процедуру можно повторить несколько раз, выбирая каждый раз новую начальную точку, и выбрать в качестве финального решения «чемпиона», т. е. решение, характеризующее наилучшими итоговыми характеристиками обслуживания. Однако, как показали практические расчеты, локальные оптимумы, полученные при различных начальных условиях, весьма близки по значению целевой функции.

10.3.3. Построение зон обслуживания с несколькими передвижными мастерскими

Рассмотрим ситуацию, изображенную на рис. 10.4, когда некоторая территория уже покрыта зонами обслуживания с единичными ПРМ. Представим себе, что зона с максимально допустимым

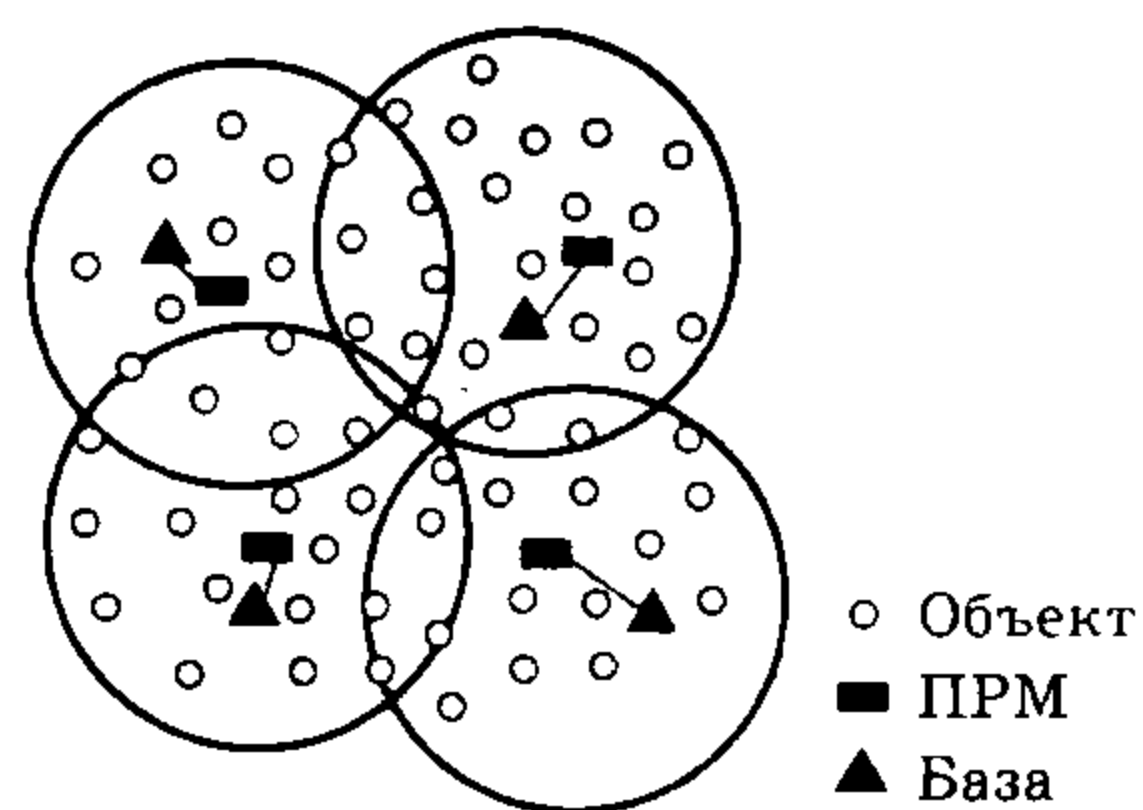


Рис. 10.4

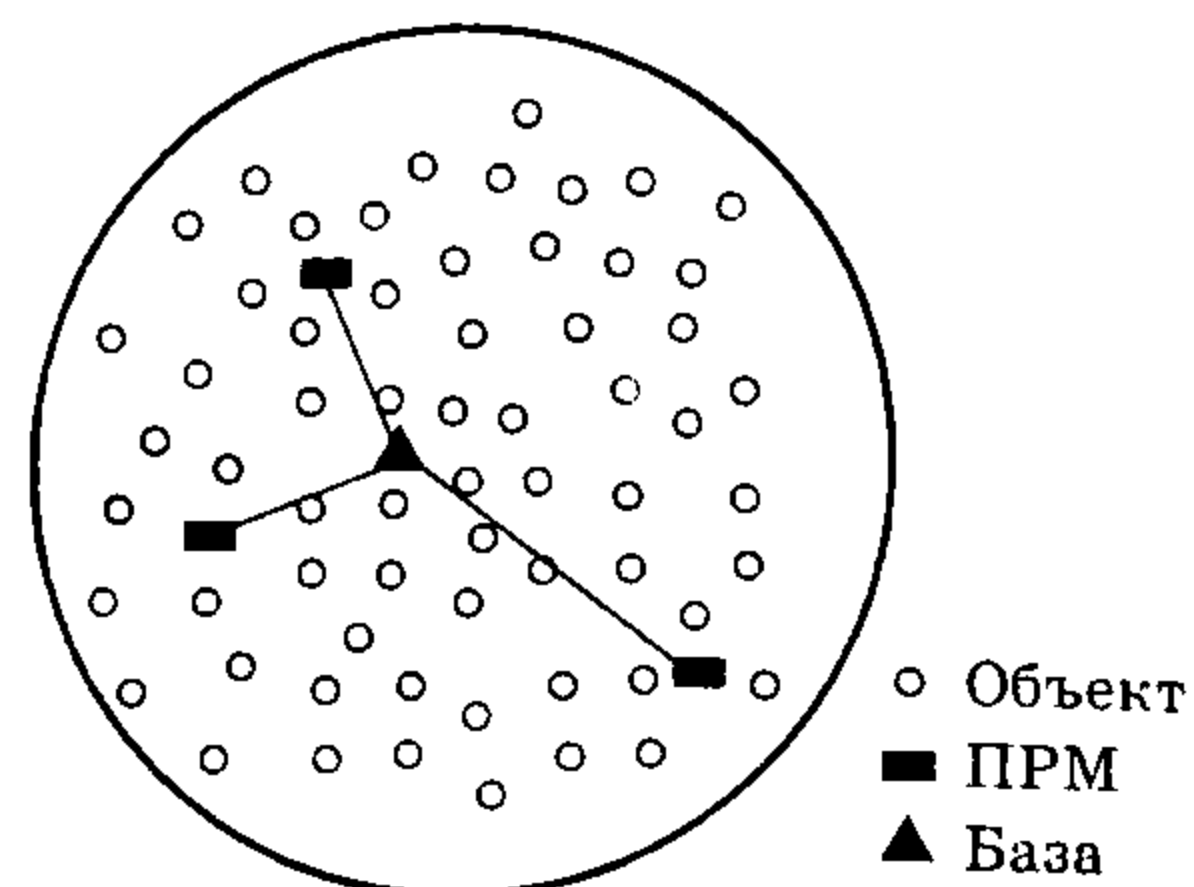


Рис. 10.5

радиусом может покрыть все эти зоны (рис. 10.5). В этом случае может оказаться, что общее число ПРМ удастся снизить, не ухудшив характеристики обслуживания.

10.3.4. Сравнение зон с индивидуальным и групповым обслуживанием

Проведем сравнение нескольких автономных зон обслуживания с индивидуальными ПРМ с одной агрегированной зоной той же общей площадью, где для обслуживания клиентов используется несколько ПРМ (рис. 10.6).

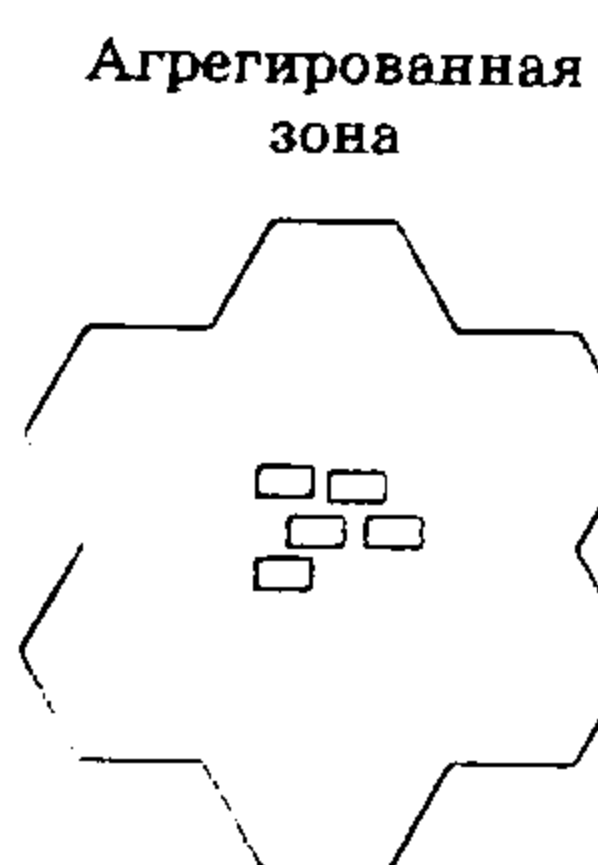
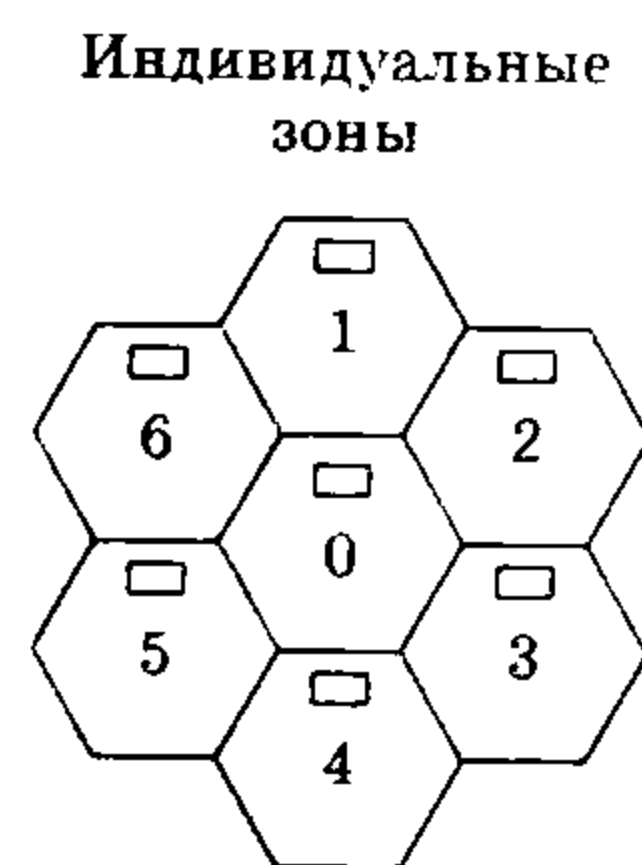


Рис. 10.6

Из теории массового обслуживания известно, что вариант с групповым обслуживанием является более эффективным. Это ясно и из общих соображений, поскольку многоканальная система обслуживания имеет большее число степеней свободы по сравнению с совокупностью одноканальных систем с закрепленной нагрузкой.

Результаты количественного сравнения среднего времени обслуживания приведены в табл. 10.1. Рассмотрены случаи высокой загрузки, т. е. те, в которых агрегирование зон наиболее рационально.

Естественно, что для агрегированной зоны обслуживания, состоящей из семи единичных зон, суммарная нагрузка была взята в семь раз больше. Из сравнения видно, что при $\rho = 0.7$ использование агрегированной зоны снижает время ожидания почти на 35% (2.5 ч вместо 3.7 ч)

Таблица 10.1

Радиус зоны = 40 км Среднее время в пути — 1 ч		Радиус зоны = 120 км Среднее время в пути — 3 ч		Число ПРМ
ρ	Среднее время обслуживания	ρ	Среднее время обслуживания	
0.7	3.7 ч	4.9	2.5 ч	6
0.8	5.1 ч	5.6	1.8 ч	7
0.9	9.1 ч	6.3	5.1 ч	7
0.99	—	≈ 7	3.2 ч	8

и при этом можно обойтись меньшим числом ПРМ. Для $\rho = 0.8$ нет выигрыша в числе ПРМ, однако время ожидания практически в три раза меньше, чем в случае одиночных зон. При более высокой интенсивности загрузки выигрыш становится еще более ощутимым.

§ 10.4. ПРИМЕР ПОСТРОЕНИЯ ЗОН ОБСЛУЖИВАНИЯ ПЕРЕДВИЖНЫМИ РЕМОНТНЫМИ МАСТЕРСКИМИ

Рассмотрим теперь практическое приложение предложенного метода к нахождению зон обслуживания для одной из систем коммерческой сети передачи цифровой информации*. Приведенный пример относится к построению зон обслуживания в южной части американского штата Флорида. В качестве минимальных территориальных единиц выбраны южные графства, затемненные на рис. 10.7.

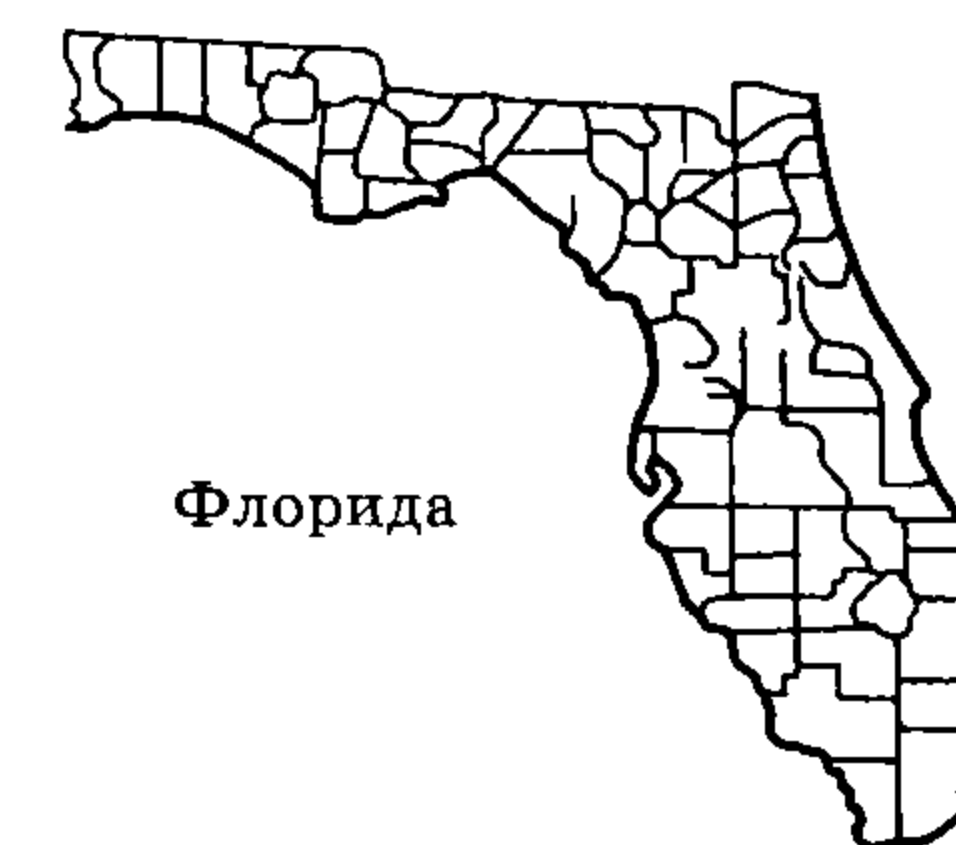


Рис. 10.7

10.4.1. Входные данные

Для построения зон обслуживания были использованы имеющиеся полевые статистические данные, а также данные о площадях различных графств Флориды, взятые из Атласа США; исходные сведения представлены в табл. 10.2.

* Проект был выполнен автором для американской компании Hughes Network Systems, Inc. (Germantown, Maryland, США).

Таблица 10.2

Номер графства	Название графства	Площадь (кв. мили)	Частота отказов (в час)
1	Dade	1955	0.047
2	Monroe	1034	0.005
3	Broward	1211	0.054
4	Collier	1994	0.010
6	Hendry	1163	0.001
5	Palm Beach	1993	0.056
7	Martin	555	0.005

Были выдвинуты следующие требования к качеству обслуживания: время ожидания обслуживания не должно превышать двух часов, а время в пути — трех часов.

10.4.2. Построение зон обслуживания единичной ремонтной мастерской

ПОСТРОЕНИЕ ПЕРВОЙ ЗОНЫ

■ **Шаг 1.** Первым выберем графство № 1, расположенное на юго-востоке полуострова Флорида, выделенное на карте рис. 10.8.

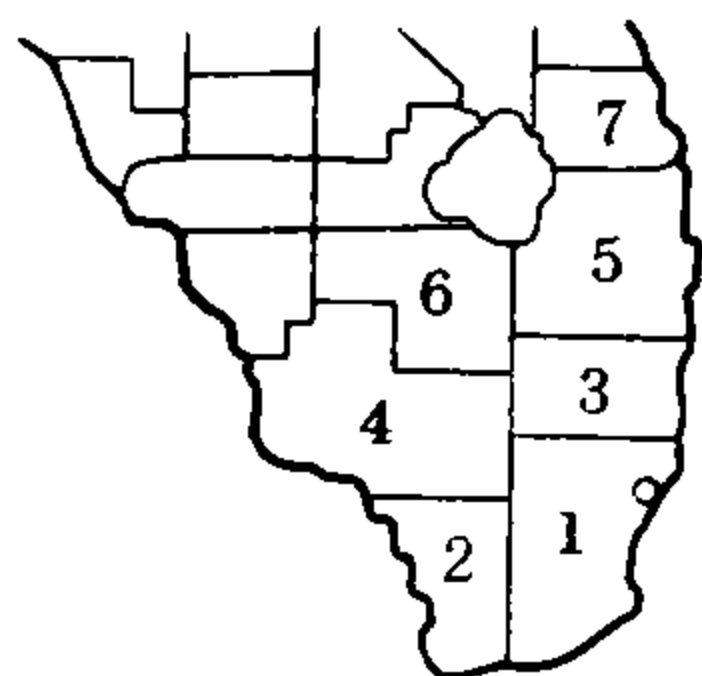


Рис. 10.8

База ПРМ размещается в точке с наивысшей концентрацией обслуживаемых объектов, в данном случае в г. Майами, обозначенном светлым кружочком.

Результаты расчетов помещены в табл. 10.3. Для вычислений была использована программа на *Microsoft Excel*, разработанная специально для решения этой задачи. Вычисленная величина времени ожидания равна 0.5 ч.

Таблица 10.3

Номер графства	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
Результат		1	0.5		0.67	0.17	24.9

■ **Шаг 2.** Расширим зону обслуживания, поскольку все характеристики обслуживания более чем удовлетворяют требованиям. В качестве следующего соседнего графства выберем графство № 2. На рис. 10.9 вновь добавленное графство закрашено темно-серым цветом, а предыдущее графство № 1 закрашено светло-серым цветом.

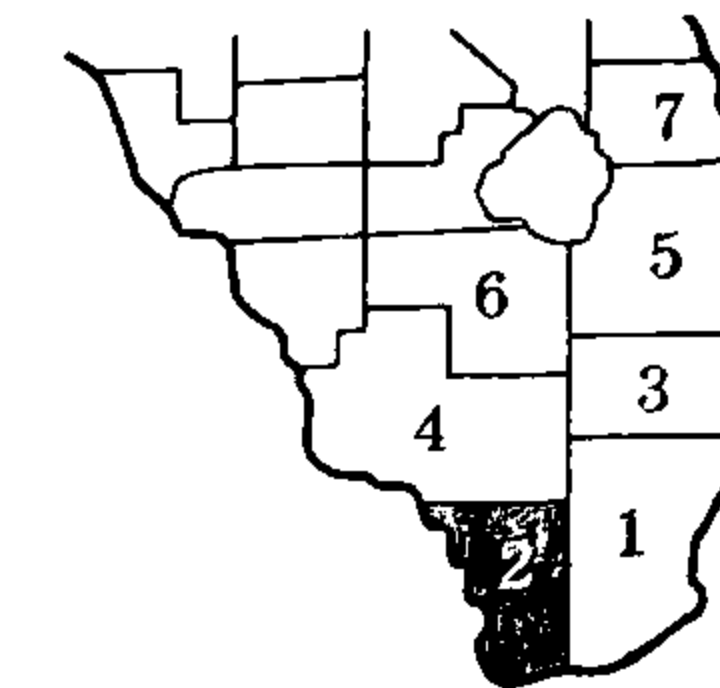


Рис. 10.9

Результаты расчетов помещены в табл. 10.4. Время ожидания равно 0.8 ч.

■ **Шаг 3.** Добавим очередное соседнее графство — № 3, выделенное темно-серым цветом на рис. 10.10.

Результаты расчетов приведены в табл. 10.5. В соответствии с требованиями время ожидания не должно превышать двух часов, однако в данном случае это требование нарушено. Следовательно, формирование такой зоны обслуживания исключено. Попробуем на следующем шаге вместо графства № 3 включить в рассмотрение графство № 4.

■ **Шаг 3-а (вторая попытка шага 3).** Добавлено графство № 4, выделенное темно-серым цветом на рис. 10.11.

Результаты расчетов помещены в табл. 10.6.

Таблица 10.4

Номер графства	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
Результат	0.052	1	0.8	2989	0.82	0.18	30.8

Таблица 10.5

Номер графства	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
3	0.054			1211			
Результат	0.106	1	2.3	2989	0.98	0.37	36.6

Таблица 10.6

	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
4	0.010			1994			
Результат	0.062	1	1.1	4983	1.06	0.22	39.8

Таблица 10.7

	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
4	0.010			1994			
6	0.001			1163			
Результат	0.062	1	1.3	6146	1.18	0.22	44.2

Таблица 10.8

	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
3	0.054			1211			
5	0.056			1993			
Результат	0.110	1	2.1	3204	0.85	0.39	31.9

Время ожидания находится в допустимых пределах, следовательно, делаем попытку расширения зоны обслуживания.

Шаг 4. Добавим графство № 6 в северной части зоны обслуживания, выделенное темно-серым цветом на рис. 10.12. Результаты вычисления сведены в табл. 10.7.

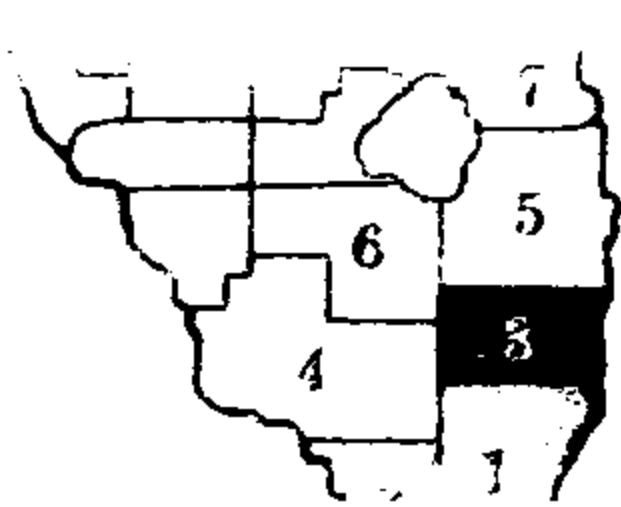


Рис. 10.11

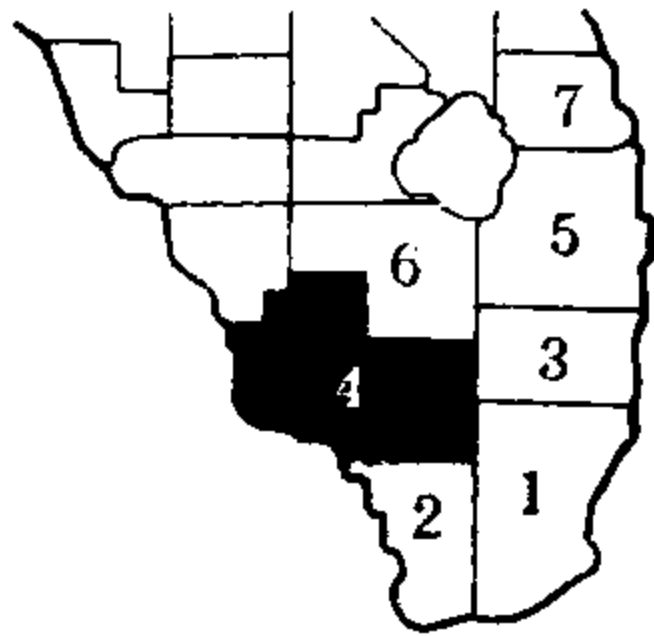


Рис. 10.12

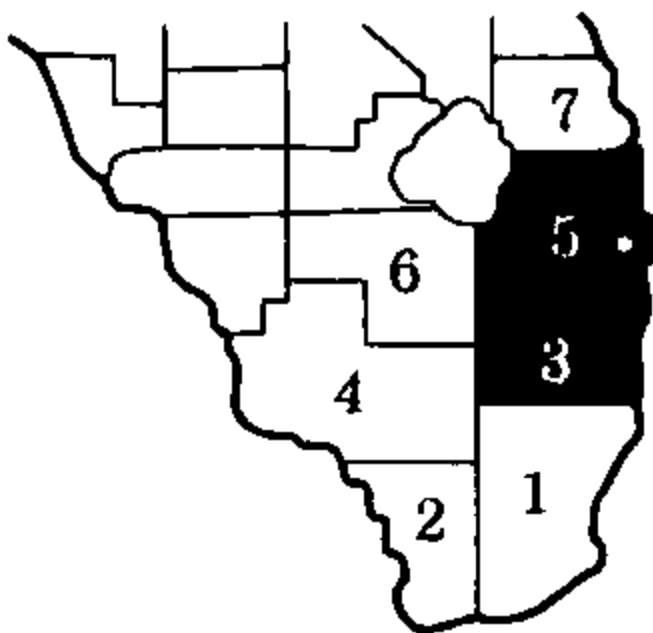
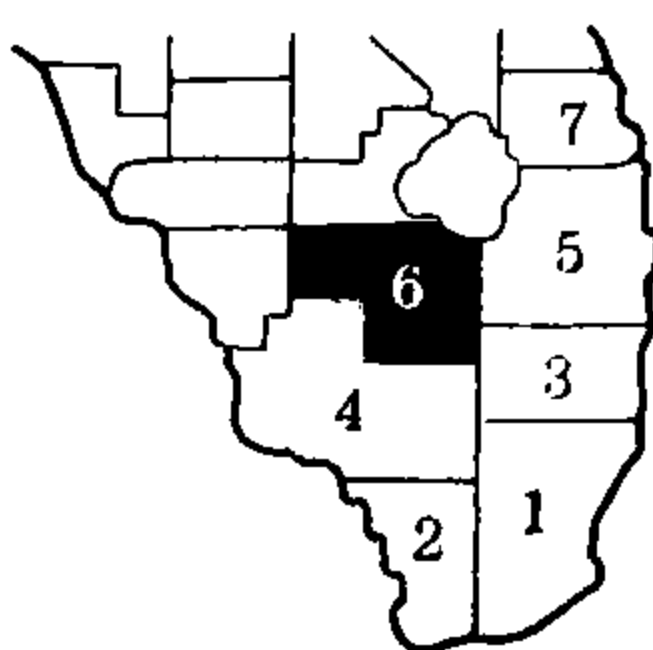


Рис. 10.13

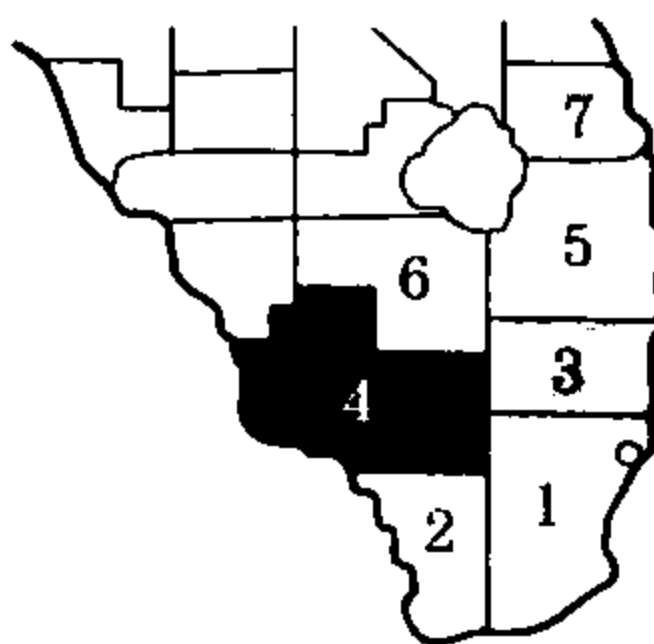


Рис. 10.14

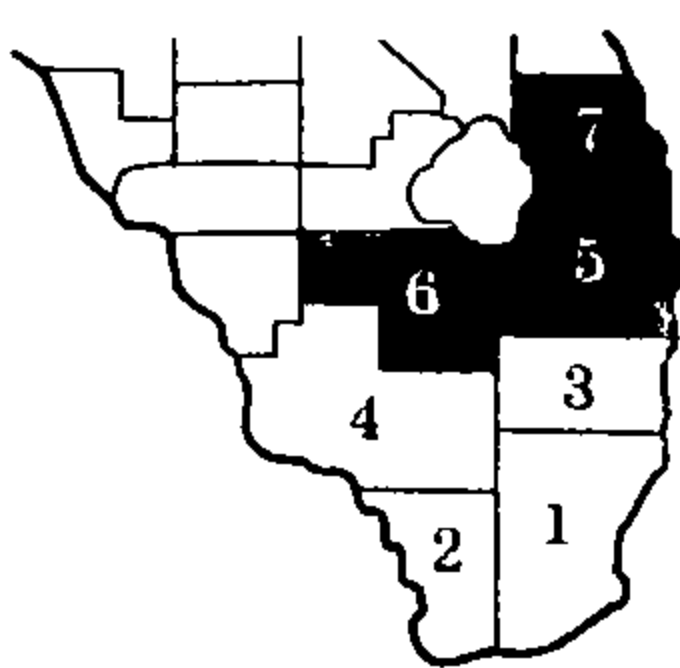


Рис. 10.15

Хотя имеется возможность включить еще какое-нибудь соседнее графство в данную зону (кроме графства № 3), мы остановимся на этом шаге, потому что дальнейшее ее разрастание приводит к ухудшению формы зоны. Собственно, такого рода качественные, плохо формализуемые аргументы и приводят к интерактивному решению задачи. Итак, построение данной зоны закончено.

ПОСТРОЕНИЕ ВТОРОЙ ЗОНЫ ОБСЛУЖИВАНИЯ. Пропуская промежуточные вычисления, начнем этот шаг с графства № 3, добавив к нему сразу же графство № 5. Базу ПРМ разместим в г. Уэст-Палм-Бич с наиболее плотным размещением обслуживаемых объектов, обозначенном светлым кружочком на рис. 10.13. Результаты вычислений приведены в табл. 10.8.

Значение времени ожидания слегка превышает требуемое, однако такая погрешность может считаться допустимой. К тому же первая из построенных зон обслуживания недогружена и может в случае надобности оказать необходимую поддержку. Такого рода качественные неформальные рассуждения чрезвычайно важны при решении практических задач.

10.4.3. Пример построения зоны обслуживания с двумя передвижными ремонтными мастерскими

В предыдущем разделе мы рассматривали построение зон с единичной ПРМ. На шаге 3 (рис. 10.10 и табл. 10.5) было получено время ожидания больше допустимого. Попробуем теперь использовать две ПРМ в этой зоне действия и даже расширить зону действия. Фактически ограничением в данном случае является максимально допустимый радиус действия ПРМ. Сразу же даже без предварительных расчетов добавим еще одно графство № 4, выделенное темно-серым цветом на рис. 10.14. Базу ПРМ сохраним по-прежнему в Майами (графство № 1), обозначенном белым кружочком. Результаты вычислений среднего времени ожидания и времени в пути даны в табл. 10.9.

Из табл. 10.9 видно, что среднее время ожидания при двух ПРМ становится заметно меньше, чем ранее — сравним с табл. 10.3, — даже при существенном увеличении нагрузки на систему. Далее, добавив еще три графта на севере зоны (№ 5, 6, 7), пропуская промежуточные шаги, приходим к зоне, представленной на рис. 10.15. Уменьшение среднего времени ожидания объясняется тем, что ожидание обслуживания происходит только в случае, когда обе ПРМ уже заняты обслуживанием текущих заявок. Результаты расчетов приведены в табл. 10.10.

Решение улучшится, если у агрегированной зоны создать две базы ПРМ, размещенные в Майами и Уэст-Палм-Биче.

Приведенные результаты показывают применимость данного метода для решения практических задач. Метод является существенно интерактивным, поскольку на каждом шаге процесса аналитик сам должен принимать решение относительно того, как расширять зону обслуживания и когда прекращать процесс, переходя к следующему этапу расчета.

Таблица 10.9

Графт	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
3	0.054			1211			
4	0.010			1994			
Результат	0.116	2	0.17	6194	1.97	0.40	41.4

Таблица 10.10

Номер графта	Частота отказов	Число ПРМ	Время ожидания	Площадь	Время в пути	ρ	Радиус
1	0.047			1955			
2	0.005			1034			
3	0.054			1211			
4	0.010			1994			
5	0.001			1163			
6	0.056			1993			
7	0.005			555			
Результат	0.178	2	0.57	9905	2.5	0.61	56.2

Приложение 1

Основные понятия и термины теории вероятностей

Ниже в конспективной форме излагаются основные сведения из теории вероятностей, необходимые для выбора и анализа показателей надежности.

- ★ **Распределение вероятности** $F(x)$ случайной величины X определяется как

$$F(x) = \Pr\{X \leq x\}. \quad (\text{п } 1.1)$$

Если случайная величина X есть наработка на отказ, то $F(x)$ имеет смысл вероятности отказа. Дополнительная величина

$$P(x) = 1 - F(x). \quad (\text{п } 1.2)$$

соответствует вероятности безотказной работы.

- **Функция плотности распределения вероятности** непрерывной случайной величины $f(x)$ определяется как

$$f(x) = \frac{d}{dx} F(x). \quad (\text{п } 1.3)$$

Очевидно, что можно записать это же соотношение в форме

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (\text{п } 1.4)$$

- **Начальным моментом** k -го порядка распределения вероятности $F(x)$ называется величина

$$M_k = \int_{-\infty}^{\infty} x^k f(x) dx = \int_{-\infty}^{\infty} x^k dF(x). \quad (\text{п } 1.5)$$

- **Начальный момент** 1-го порядка

$$M_1 = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} x dF(x) \quad (\text{п } 1.6)$$

называется **математическим ожиданием** или **средним** случайной величины X .

Математическое ожидание по своему формальному определению совпадает с определением центра тяжести неоднородного тонкого стержня с распределением массы по длине в соответствии с $f(x)$ (рис. П 1.1).

Для непрерывных неотрицательных случайных величин, какими, в частности, являются все наработки на отказ, можно записать следующее выражение для среднего:

$$M_1 = \int_0^{\infty} P(t) dt. \quad (\text{П } 1.7)$$

Действительно,

$$M_1 = E\{X\} = \int_0^{\infty} t f(t) dt = \int_0^{\infty} t dF(t) = - \int_0^{\infty} t dP(t) = -tP(t) \Big|_0^{\infty} + \int_0^{\infty} P(t) dt,$$

где $-tP(t) \Big|_0^{\infty} = 0$ для всех собственных функций распределения.

■ Математическое ожидание суммы случайных величин равно сумме математических ожиданий этих случайных величин, т. е.

$$E\{X_1 + X_2 + \dots + X_n\} = E\{X_1\} + E\{X_2\} + \dots + E\{X_n\}. \quad (\text{П } 1.8)$$

■ Центральным моментом k -го порядка распределения вероятностей $F(x)$ называется величина

$$M_{(k)} = \int_{-\infty}^{\infty} (x - M_1)^k f(x) dx = \int_{-\infty}^{\infty} (x - M_1)^k dF(x). \quad (\text{П } 1.9)$$

■ Дисперсией или центральным моментом второго порядка называется величина

$$\sigma^2 = M_{(2)} = \int_{-\infty}^{\infty} (x - M_1)^2 f(x) dx = \int_{-\infty}^{\infty} (x - M_1)^2 dF(x). \quad (\text{П } 1.10)$$

■ Проведя простые преобразования, получаем

$$\int_{-\infty}^{\infty} (x - M_1)^2 f(x) dx = \int_{-\infty}^{\infty} (x^2 - 2M_1x + M_1^2) f(x) dx = M_2 - M_1^2,$$

или окончательно

$$\sigma^2 = M_2 - M_1^2. \quad (\text{П } 1.11)$$

Величина σ называется средним квадратическим отклонением.

По формальному определению дисперсия аналогична моменту инерции. На самом деле дисперсия характеризует разброс «вероятностной массы» относительно «центра тяжести» распределения. На рис. П 1.2 представлены случаи распределений с различными дисперсиями.

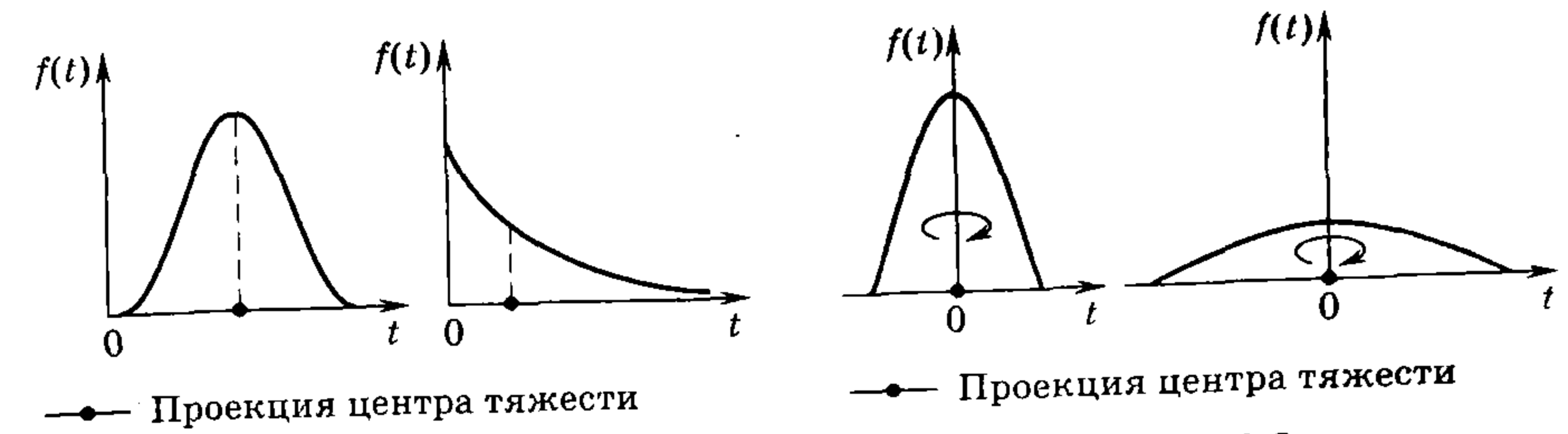


Рис. П 1.1

Рис. П 1.2

Для независимых случайных величин X_1, X_2, \dots, X_n дисперсия $\sigma^2(X)$ суммы этих случайных величин равна сумме дисперсий этих случайных величин $\sigma^2(X_i)$, т. е.

$$\sigma^2\{X_1 + X_2 + \dots + X_n\} = \sigma^2\{X_1\} + \sigma^2\{X_2\} + \dots + \sigma^2\{X_n\}. \quad (\text{П } 1.12)$$

Приложение 2

Основные распределения, используемые в теории надежности

Показатели надежности представляют собой инженерные интерпретации тех или иных характеристик распределений случайных величин или случайных процессов. Здесь приведены основные распределения вероятностей, используемые в теории надежности.

П 2.1. РАСПРЕДЕЛЕНИЕ БЕРНУЛЛИ

Пусть возможны две взаимоисключающие друг друга реализации события X : успех и отказ. Пусть успех реализуется с вероятностью p , а отказ — с вероятностью $q = 1 - p$. Присвоим событию успех значение 1, а событию отказ значение 0, т. е. определим случайную величину X как:

$$X = \begin{cases} 1, & \text{если наблюдается успех,} \\ 0, & \text{если наблюдается отказ.} \end{cases}$$

Функция распределения случайной величины X может быть записана в виде

$$\Pr \{X = x\} = \begin{cases} p, & \text{если } x = 1, \\ q, & \text{если } x = 0, \end{cases} \quad (\text{П 2.1})$$

или в компактной форме

$$\{X = x\} = p^x q^{x-1}. \quad (\text{П 2.2})$$

Распределение (2.1) называется распределением Бернулли, а случайная величина X — бернуллиевой случайной величиной. Среднее значение этого распределения равно

$$M_1 = 1 \cdot p + 0 \cdot q = p, \quad (\text{П 2.3})$$

и второй начальный момент равен

$$M_2 = 1^2 \cdot p + 0^2 \cdot q = p. \quad (\text{П 2.4})$$

Дисперсия вычисляется по формуле (П 1.11)

$$\sigma^2\{X\} = M^2 - (M_1)^2 = p - p^2 = pq. \quad (\text{П 2.5})$$

Распределение Бернулли применяется при описании явлений с двумя возможными исходами, например такими, как бросание монеты.

П 2.2. ГЕОМЕТРИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ

Рассмотрим последовательность независимых испытаний Бернулли. Пусть X есть случайная величина, равная числу последовательных успехов до первого отказа. Такая случайная величина X имеет геометрическое распределение:

$$\{X = k\} = p^k q. \quad (\text{П 2.6})$$

Кумулятивная функция этого распределения есть

$$\begin{aligned} \Pr\{X \leq x\} &= q \sum_{k=0}^x p^k = 1 - q \sum_{k=x+1}^{\infty} p^k = \\ &= 1 - qp^{x+1} \sum_{k=0}^{\infty} p^k = 1 - \frac{qp^{x+1}}{1-p} = 1 - p^{x+1}. \end{aligned} \quad (\text{П 2.7})$$

Заметим, что эта вероятность может быть найдена и на основании других рассуждений. Вероятность $P\{X > x\}$ соответствует событию, заключающемуся в том, что произойдут первые $(x + 1)$ успехов, и не важно, что произойдет после этого. Это событие происходит с вероятностью p^{x+1} , т. е.

$$P\{X > x\} = p^{x+1}, \quad (\text{П 2.8})$$

а следовательно, дополнительная вероятность равна

$$P\{X \leq x\} = 1 - p^{x+1}.$$

Среднее значение случайной величины X находится обычным способом:

$$M_1 = \sum_{x=0}^{\infty} xp^x q. \quad (\text{П 2.9})$$

Вычисления приводят нас к следующему:

$$\sum_{x=0}^{\infty} xp^x q = pq \cdot \frac{d}{dp} \sum_{x=0}^{\infty} p^x = pq \cdot \frac{d}{dp} \left(\frac{1}{1-p} \right) = \frac{p}{q}. \quad (\text{П 2.10})$$

Геометрическим распределением описывается, например, длина серии успешных исходов на испытаниях образцов, а также число последовательных удачных срабатываний реле или переключателей.

П 2.3. ЭКСПОНЕНЦИАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Рассмотрим геометрическое распределение, у которого вероятность успеха, p , близка к единице, или, что то же самое, значение $q = 1 - p \ll 1$. Вероятность того, что случайная величина X , равная числу успехов в непрерывной последовательности, будет больше заданного n , равна

$$\Pr\{X \geq n\} = p^n = (1 - q)^n. \quad (\text{П 2.11})$$

Пусть каждое испытание Бернулли длится время Δ . Вероятность отказа в течение этого времени можно записать в виде $q = \lambda\Delta$, где λ есть некоторая константа, смысл которой будет пояснен ниже. Допустим, что проведено достаточно большое число опытов n . Пусть n испытаний Бернулли длятся в течение времени $t = n\Delta$. Тогда формулу (П 2.11) можно переписать в виде

$$P(t) = P\{X \geq n\} = (1 - \lambda\Delta)^{t/\Delta},$$

откуда, переходя к пределу при $\Delta \rightarrow 0$, получаем

$$P(t) = \exp(-\lambda t). \quad (\text{П 2.12})$$

Поскольку (П 2.12) есть функция, дополнительная к функции распределения, для функции распределения можно записать

$$F(t) = 1 - \exp(-\lambda t). \quad (\text{П 2.13})$$

Полученное распределение называется экспоненциальным распределением, а λ — параметром этого распределения.

Таким образом, экспоненциальное распределение в определенном смысле является предельным для геометрического.

Плотность экспоненциального распределения находится как

$$f(t) = \frac{d}{dt} [1 - \exp(-\lambda t)] = \lambda \exp(-\lambda t). \quad (\text{П 2.14})$$

Среднее для экспоненциального распределения можно легко найти путем прямого интегрирования

$$M_1 = \int_0^{\infty} t \lambda \exp(-\lambda t) dt = \frac{1}{\lambda}. \quad (\text{П 2.15})$$

Второй начальный момент находится как

$$M_2 = \int_0^{\infty} \lambda t^2 \exp(-\lambda t) dt = \frac{2}{\lambda^2}. \quad (\text{П 2.16})$$

Используя определение дисперсии (П 1.7), получаем из (П 2.15) и (П 2.16) дисперсию экспоненциального распределения:

$$\sigma^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \quad (\text{П 2.17})$$

Экспоненциальное распределение, как и геометрическое, обладает марковским свойством. В частности, это означает, что если $X > x$, т. е. до момента x отказ не произошел, то

$$P(x, t + x | X > x) = P(0, t) = \exp(-\lambda t). \quad (\text{П 2.18})$$

Иными словами, если некоторый объект характеризуется экспоненциальным временем работы до отказа, то объект, проработавший произвольное время, но не отказавший к данному моменту, по своим характеристикам надежности будет неотличим от совершенно нового объекта.

Экспоненциальное распределение наиболее часто используется в расчетах надежности, поскольку обычно бывает известна только средняя наработка на отказ или обратная ей величина — интенсивность отказов, т. е. использование однопараметрического распределения естественно. Второй немаловажной причиной использования экспоненциального распределения в инженерных расчетах, хотя, вероятно, и не всегда оправданной*, является возможность получать удобные расчетные формулы.

П 2.4. ГИПЕРЭКСПОНЕНЦИАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Иногда встречаются ситуации, когда время безотказной работы или время восстановления имеют гиперэкспоненциальное распределение:

$$P(t) = \sum p_k \exp(-\lambda_k t). \quad (\text{П 2.19})$$

Для такого распределения среднее легко находится как средневзвешенное значение средних исходных экспоненциальных распределений:

$$M_1 = \sum_{k=1}^n \frac{p_k}{\lambda_k}. \quad (\text{П 2.20})$$

Данное распределение имеет убывающую интенсивность отказов:

$$\lambda(t) = \frac{\sum_{k=1}^n p_k \lambda_k \exp(-\lambda_k t)}{\sum_{k=1}^n p_k \exp(-\lambda_k t)}, \quad (\text{П 2.21})$$

* Иногда, правда, использование экспоненциального распределения без должного обоснования напоминает поиск потерянных на улице ключей ночью под фонарем лишь потому, что под ним светло!

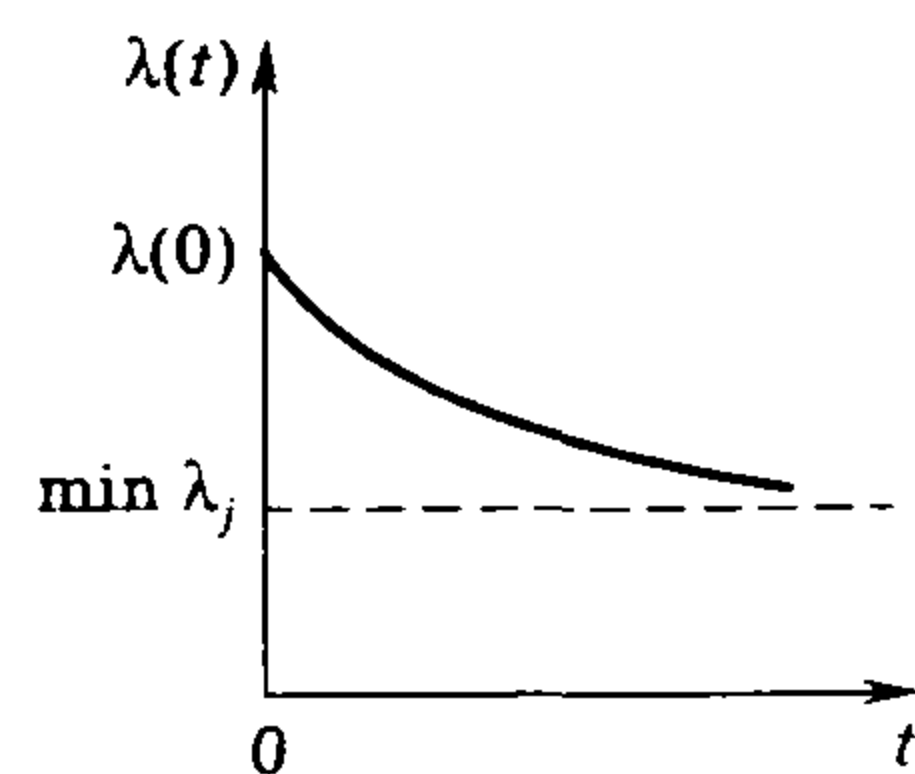


Рис. П 2.1

причем при $t = 0$

$$\lambda(0) = \frac{\sum_{k=1}^n p_k \lambda_k}{\sum_{k=1}^n p_k} = \sum_{k=1}^n p_k \lambda_k, \quad (\text{П 2.22})$$

а при $t \rightarrow \infty$, как нетрудно видеть непосредственно из (П 2.21),

$$\lim_{t \rightarrow \infty} \lambda(t) = \min_{1 \leq j \leq n} \lambda_j. \quad (\text{П 2.23})$$

Очевидно, $\lambda(0) \geq \min_{1 \leq j \leq n} \lambda_j$. Таким образом, функция интенсивности отказов в этом случае имеет вид, представленный на рис. П 2.1.

Такое распределение *взвешенной смеси n экспонент* возникает в случае, когда в системе используются элементы, принадлежащие разным поставщикам или разным партиям продукции, каждая из которых сама по себе может иметь и весьма стабильные показатели надежности, но сами эти показатели существенно различаются между собой. Аналогичная ситуация возникает с временем ремонта, когда имеются ремонтные органы с различной скоростью проведения восстановительных работ. В этом случае результирующее время восстановления имеет гиперэкспоненциальное распределение.

П 2.5. БИНОМИАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Представим себе, что проводится серия n испытаний Бернулли. Какова вероятность того, что в течение этих испытаний произойдет ровно k успехов? Рассматриваемая случайная величина X представляет собой сумму n независимых бернуллиевых случайных величин, т. е.

$$X = X_1 + X_2 + \dots + X_n = \sum_{k=1}^n X_k, \quad (\text{П 2.24})$$

где X_k есть бернуллиева случайная величина, соответствующая k -му испытанию.

Случайная величина X , определенная таким образом, подчиняется биномиальному закону распределения. Вероятность того, что при n испытаниях Бернулли произойдет ровно k успехов, определяется как

$$P(n; k) = \binom{n}{k} p^k q^{n-k}. \quad (\text{П 2.25})$$

Воспользовавшись формулой (П 1.4) и взяв значение для среднего в распределении Бернулли (П 2.3), можно записать

$$E\{X_1 + X_2 + \dots + X_n\} = nE\{X_1\} = np. \quad (\text{П 2.26})$$

Воспользовавшись формулой (П 1.9) и имея в виду (П 2.5), можно записать для биномиального распределения

$$\sigma^2 = npq. \quad (\text{П 2.27})$$

Отметим важное свойство биномиального распределения: если производятся две последовательности испытаний Бернулли, одна объема n_1 , а другая — объема n_2 , причем обе с одинаковой вероятностью успеха p , то это эквивалентно проведению одной последовательности испытаний Бернулли объема $n = n_1 + n_2$. Конечно, этот результат интуитивно ясен. Он легко может быть и строго доказан при помощи производящей функции.

В инженерной практике биномиальное распределение встречается при решении задач статистического контроля готовой продукции, при анализе систем с нагруженным резервом и в некоторых других случаях.

П 2.6. РАСПРЕДЕЛЕНИЕ ПУАССОНА

Распределение Пуассона очень часто используется при построении различных математических моделей в теории надежности. Рассмотрим биномиальное распределение для случая, когда число испытаний Бернулли n велико, а вероятность p близка к единице. В этом случае вычисление вероятности (П 2.25) для биномиального распределения представляет определенную вычислительную трудность: приходится иметь дело с большим числом слагаемых, каждое из которых является произведением больших значений биномиальных коэффициентов на очень малые значения вероятностей. Рассмотрим предельное поведение вероятности (П 2.25), когда $n \rightarrow \infty$, $q \rightarrow 0$, а $nq = \text{const}$:

$$\binom{n}{k} q^k p^{n-k} \approx \frac{q^k}{k!} \lim [n \cdot (n-1) \cdot \dots \cdot (n-k+1)] (1-q)^{n-k}.$$

Поскольку $n \geq k$, имеем $[n \cdot (n-1) \cdot \dots \cdot (n-k+1)] \approx n^k$ и $\lim_{n \rightarrow \infty} (1-q)^{n-k} = \exp(-nq)$. В итоге получаем

$$\lim_{n \rightarrow \infty} \binom{n}{k} q^k p^{n-k} = \frac{(nq)^k}{k!} \exp(-nq). \quad (\text{П 2.28})$$

Формула (П 2.28) отражает фундаментальную в теории вероятностей теорему Пуассона. Пусть интервал $[0, t]$ разбит на n подынтервалов длиной Δ . Если мы теперь положим, что каждое испытание Бернулли занимает Δ единиц времени, а среднее число отказов в единицу времени составляет λ , то величина nq может быть заменена на λt . Теперь формулу (П 2.28) можно переписать в виде

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (\text{П 2.29})$$

Итак, если в среднем за единицу времени происходит λ отказов, то $p_k(t)$ есть вероятность того, что за время t произойдет ровно k отказов. Кумулятивная функция запишется следующим образом:

$$P_{j \leq k}(t) = \sum_{j=0}^k \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (\text{П 2.30})$$

В определенном смысле распределение Пуассона соотносится с биномиальным распределением, как экспоненциальное распределение соотносится с геометрическим.

Находим среднее число отказов на интервале t :

$$M_1 = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \lambda t \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} = \lambda t e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = \lambda t. \quad (\text{П 2.31})$$

Таким образом, величина λt представляет собой математическое ожидание числа отказов за время t :

$$M_1 = \lambda t. \quad (\text{П 2.32})$$

Это же следует из определения λ как среднего числа отказов в единицу времени.

Дисперсия числа отказов при распределении Пуассона равна значению математического ожидания, т. е.

$$\sigma^2 = \lambda t. \quad (\text{П 2.33})$$

Это легко проверить и непосредственно, найдя второй момент распределения Пуассона:

$$\begin{aligned} M_2 &= \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \lambda t \sum_{k=0}^{\infty} (k-1+1) \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} = \\ &= \lambda t \sum_{k=0}^{\infty} (k-1) \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} + \lambda t \sum_{k=0}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} = (\lambda t)^2 + \lambda t. \end{aligned} \quad (\text{П 2.34})$$

Отсюда получаем

$$\sigma^2 = M_2 - (M_1)^2 = \lambda t. \quad (\text{П 2.35})$$

Распределение Пуассона удобно использовать для приближенных вычислений биномиальных вероятностей, когда число испытаний велико, а ожидаемое число отказов мало.

В теории надежности распределение Пуассона используется во многих прикладных задачах, в том числе в задачах, связанных с расчетом необходимого числа запасных элементов.

П 2.7. РАСПРЕДЕЛЕНИЕ ЭРЛАНГА

Распределением Эрланга k -го порядка называется распределение суммы k идентичных* экспоненциально распределенных случайных величин.

Обозначим плотность распределения этой случайной величины через $e_k(t)$. Распределение Эрланга 1-го порядка совпадает с экспоненциальным, т. е. его плотность выражается как

$$e_1(t) = \exp(-\lambda t). \quad (\text{П 2.36})$$

Рекуррентным образом можно получить выражение для плотности распределения Эрланга k -го порядка, $e_k(t)$:

$$e_k(t) = \int_0^t e_{k-1}(\lambda y) \lambda \exp[-(t-y)\lambda] dy = \lambda \frac{(\lambda t)^{k-1}}{(k-1)!} \exp(-\lambda t). \quad (\text{П 2.37})$$

Поскольку распределение Эрланга k -го порядка есть распределение суммы экспоненциальных величин, это позволяет с использованием (П 1.9) и (П 1.13) записать формулы для среднего и дисперсии:

$$M_1 = \frac{k}{\lambda}, \quad (\text{П 2.38})$$

$$\sigma^2 = \frac{k}{\lambda^2}. \quad (\text{П 2.39})$$

Из выражения (П 2.37) следует, что распределение Эрланга тесно связано с распределением Пуассона.

Функция интенсивности отказов для распределения Эрланга k -го порядка, которую нетрудно получить, используя (П 2.37), имеет вид

$$\lambda_k(t) = \frac{\lambda \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t}}{1 - \int_0^t \lambda \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} dt}. \quad (\text{П 2.40})$$

* Иногда рассматривается обобщенное распределение Эрланга, где слагаемые случайные величины могут иметь различные экспоненциальные распределения, каждое со своим параметром.

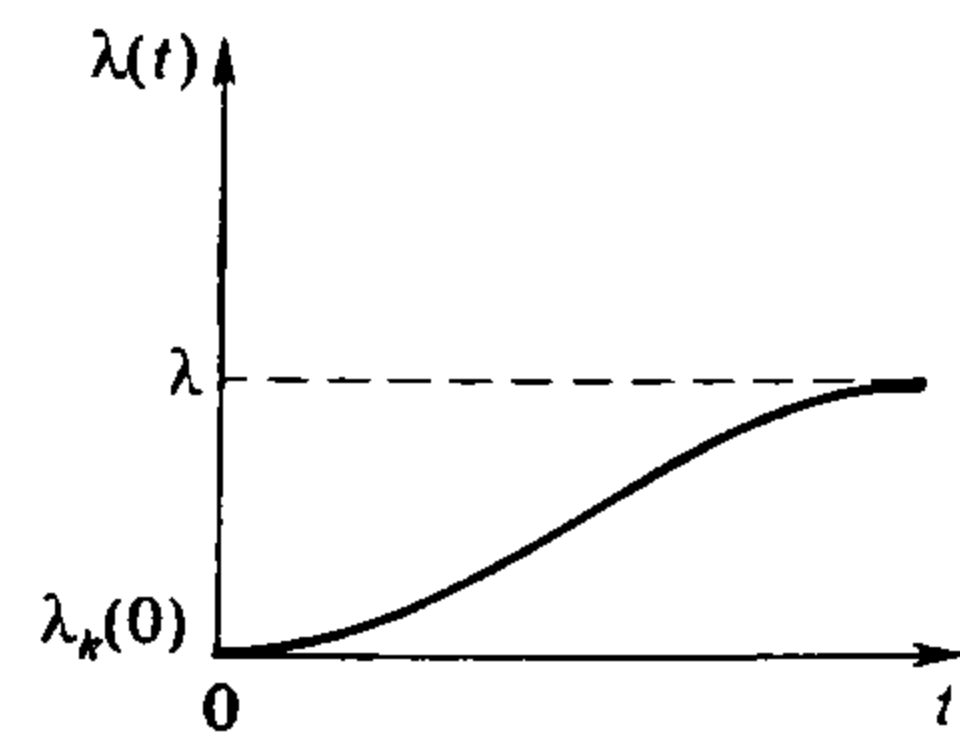


Рис. П 2.2

Качественный анализ этой функции показывает, что $\lambda_k(0) = 0$ в начальный момент времени, и $\lambda_k(t) \rightarrow \lambda$ при $t \rightarrow \infty$ (см. рис. П 2.2).

Распределение Эрланга используется в теории надежности для описания различных многофазовых процессов. Например, случайное время работы невосстанавливаемой системы с *холодным* резервом имеет распределение Эрланга; время восстановления может состоять из нескольких процедур, каждая из которых случайна и распределена экспоненциально, тогда его распределение также имеет распределение Эрланга.

П 2.8. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Нормальное распределение, часто называемое распределением Гаусса, играет важную роль во многих вероятностных приложениях. Центральная предельная теорема теории вероятностей гласит, что сумма независимых случайных величин в пределе имеет нормальное распределение (при достаточно слабых условиях). Плотность нормального распределения, определенная на всей действительной оси, имеет вид

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right), \quad (\text{П } 2.41)$$

где параметры a и σ^2 есть математическое ожидание и дисперсия соответственно. Плотность нормированного и центрированного нормального распределения, т. е. распределения с $a = 0$ и $\sigma^2 = 1$:

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), \quad (\text{П } 2.42)$$

табулирована. Интегральная (кумулятивная) функция распределения табулирована в нескольких формах:

✓ как стандартная функция распределения

$$F(t) = \int_{-\infty}^t f(x)dx, \quad (\text{П } 2.43)$$

✓ в виде функции Лапласа

$$\Phi(t) = \int_{-t}^t f(x)dx. \quad (\text{П } 2.44)$$

В теории надежности мы имеем дело с неотрицательными случайными величинами, имеющими смысл времени (время наработки на отказ, длительность замены и т. п.). В то же время нормальное распределение определено на всей числовой оси $(-\infty, \infty)$. В связи с этим в задачах надежности приходится рассматривать усеченное слева нормальное распределение (рис. П 2.3). Для усеченного нормального распределения плотность должна быть перенормирована таким образом, чтобы площадь под кривой от 0 до ∞ была бы равна единице.



Рис. П 2.3

Выражение для плотности усеченного нормального распределения таково:

$$\begin{aligned} f_{\text{усеч}}(t) &= \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right)}{\frac{1}{\sigma\sqrt{2\pi}} \int_0^{\infty} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right) dy} = \\ &= \frac{1}{1 - F(0)} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right). \end{aligned} \quad (\text{П } 2.45)$$

В практических задачах надежности, когда в качестве случайной величины рассматривается время работы до отказа, обычно $\sigma/a \ll 1$, а посему знаменатель в (П 2.45) очень близок к единице, т. е. фактически можно пользоваться неусеченным нормальным распределением.

Нормальное распределение широко используется для аппроксимации биномиального распределения и распределения Пуассона, когда их математические ожидания велики. Эта аппроксимация основывается на центральной предельной теореме в форме теоремы Муавра — Лапласа. Для биномиального распределения используется аппроксимирующее нормальное распределение со средним $a = np$ и дисперсией $\sigma^2 = np$, а в случае распределения Пуассона берется нормальное распределение со средним $a = \lambda t$ и дисперсией $\sigma^2 = \lambda t$.

Нормальным распределением характеризуют в некоторых случаях время безотказной работы и время ремонта.

П 2.9. РАВНОМЕРНОЕ РАСПРЕДЕЛЕНИЕ

Равномерным распределением называется распределение, определенное на интервале $[x, y]$ (рис. П 2.4) и имеющее плотность

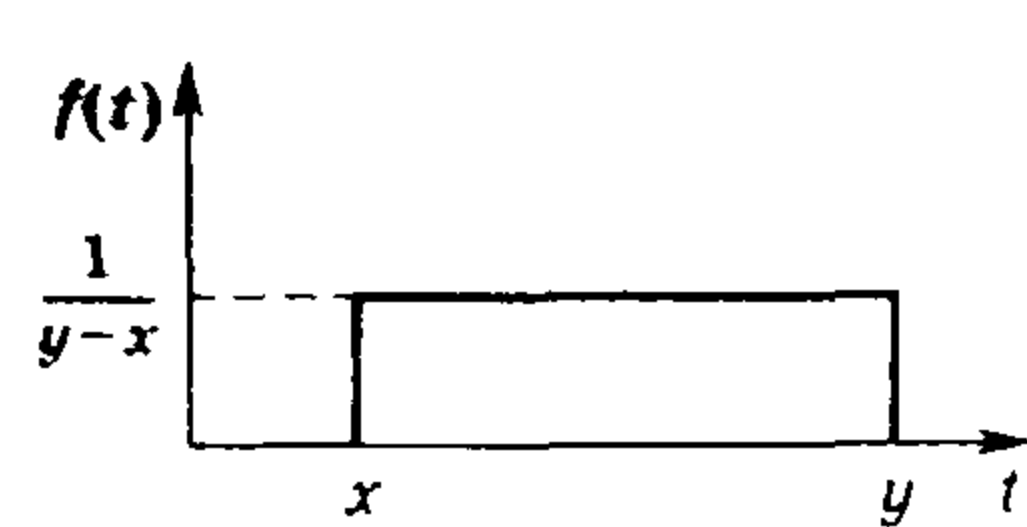


Рис. П 2.4

$$f(t) = \frac{1}{y - x}. \quad (\text{П } 2.46)$$

Среднее этого распределения, очевидно, равно

$$M_1 = \frac{y - x}{2}, \quad (\text{П } 2.47)$$

а дисперсия, как нетрудно вычислить, составляет

$$\sigma^2 = \frac{(y - x)^2}{12}. \quad (\text{П } 2.48)$$

Это распределение является базовым при моделировании методом Монте-Карло. Здесь генерация случайных величин ξ с любым заданным распределением $F(t)$ получается путем использования обратного преобразования: $\xi = F^{-1}(\psi)$, где ψ — равномерно распределенная на интервале $[0, 1]$ случайная величина.

П 2.10. ВЫРОЖДЕННОЕ РАСПРЕДЕЛЕНИЕ

Вырожденным называется распределение вида

$$F(t) = \begin{cases} 0, & \text{если } t < T, \\ 1, & \text{если } t \geq T. \end{cases} \quad (\text{П } 2.49)$$

Выражение (П 2.49) есть не что иное, как распределение постоянной величины T .

В теории надежности постоянная величина используется для описания времени переключения на резерв, а также в ряде других специальных случаев, например при получении оценок для стареющих элементов.

П 2.11. ОТНОШЕНИЯ МЕЖДУ РАСПРЕДЕЛЕНИЯМИ ВЕРОЯТНОСТЕЙ

Для разумного использования различных распределений при решении задач надежности важно «осознать» связь и взаимоотношения между различными семействами случайных величин. Попытка помочь читателю в этом направлении предпринята с помощью диаграммы, изображенной на рис. П 2.5. Опишем все «генетические» отноше-

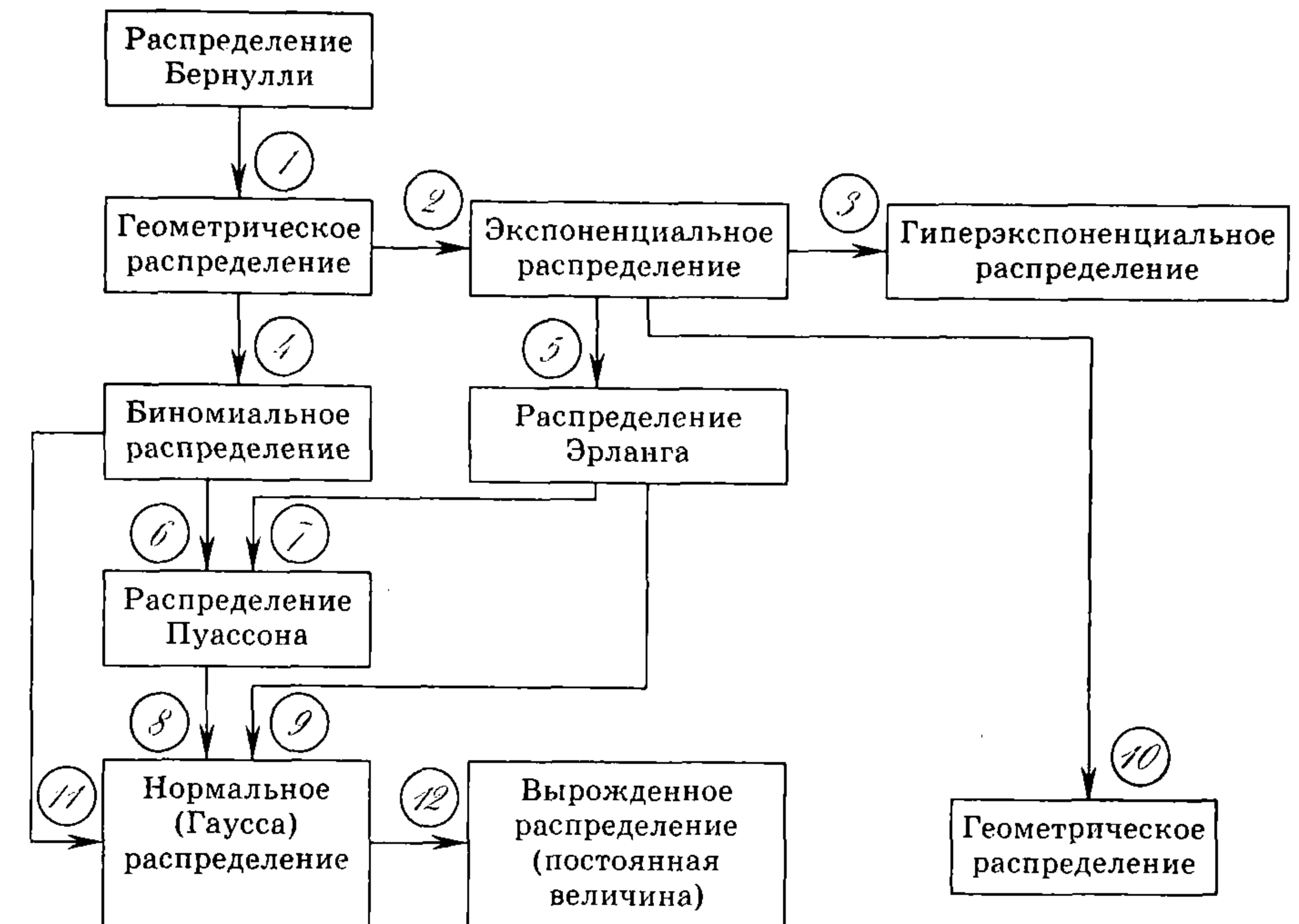


Рис. П 2.5

ния между распределениями вероятностей, представленными этой диаграммой.

- (1) Безусловно «пращуром» всех распределений вероятностей является распределение **Бернулли**: это распределение *индикатора события*, который принято обозначать через «1» для успеха в испытании и через «0» для неудачного испытания (отказа). Непрерывная серия (последовательность) однородных индикаторов (т. е. серия единиц или же нулей) приводит нас к геометрическому распределению.
- (2) Если длительность каждого испытания Бернулли устремить к нулю, а вероятности успеха устремить к единице, то при предельном переходе из геометрического распределения получается экспоненциальное.
- (3) Если образовать взвешенную сумму экспоненциальных величин, то произойдет переход от экспоненциального распределения к гиперэкспоненциальному.
- (4) Рассматривая интегральное распределение по отношению к геометрическому, получаем биномиальное распределение. В самом деле, что означает получить k успехов в n испытаниях Бернулли? Это означает найти полную

вероятность того, что сумма $(k - 1)$ геометрических случайных величин меньше n при условии, что сумма k случайных величин больше n .

- (5) Распределение Эрланга является распределением суммы фиксированного числа экспоненциально распределенных случайных величин.
- (6) Пуассоновское распределение является предельным по отношению к биномиальному.
- (7) Отношения между распределением Эрланга и распределением Пуассона напоминают отношения между геометрическим и биномиальным распределением: распределение Пуассона является интегральным по отношению к распределению Эрланга.
- (8) Нормальное (Гаусса) распределение является предельным для распределения Пуассона при больших средних значениях последнего.
- (9) Плотность нормального (Гаусса) распределения является хорошей аппроксимацией для плотности распределения Эрланга высокого порядка.
- (10) Равномерное распределение можно рассматривать как хорошую аппроксимацию экспоненциального распределения для малых значений аргумента, т. е. когда показатель экспоненты много меньше единицы.
- (11) Нормальное (Гаусса) распределение является предельным для биномиального распределения при больших средних значениях последнего.
- (12) При стремлении дисперсии нормального распределения к нулю, это распределение стремится к вырожденному со значением, равным среднему исходного нормального распределения.

П 2.12. РАСПРЕДЕЛЕНИЕ ВЕЙБУЛЛА — ГНЕДЕНКО

В заключение рассмотрим широко используемое в инженерных приложениях распределение Вейбулла — Гнеденко*. Это распределение имеет вид

$$F(t) = \begin{cases} 1 - \exp(-(\lambda t)^\beta) & \text{для } t \geq 0, \\ 0 & \text{для } t < 0. \end{cases} \quad (\text{П 2.50})$$

* Шведский инженер В. Вейбулл, изучая процесс отказов шарикоподшипников, предложил в 30-х годах прошлого столетия своеобразную модель *слабейшего звена*. Для аппроксимации эмпирических

Его плотность записывается следующим образом:

$$\psi(t) = \begin{cases} \lambda^\beta t^{\beta-1} \exp(-(\lambda t)^\beta) & \text{для } t \geq 0, \\ 0 & \text{для } t < 0. \end{cases} \quad (\text{П 2.51})$$

Это распределение характеризуется параметром масштаба λ и параметром формы β .

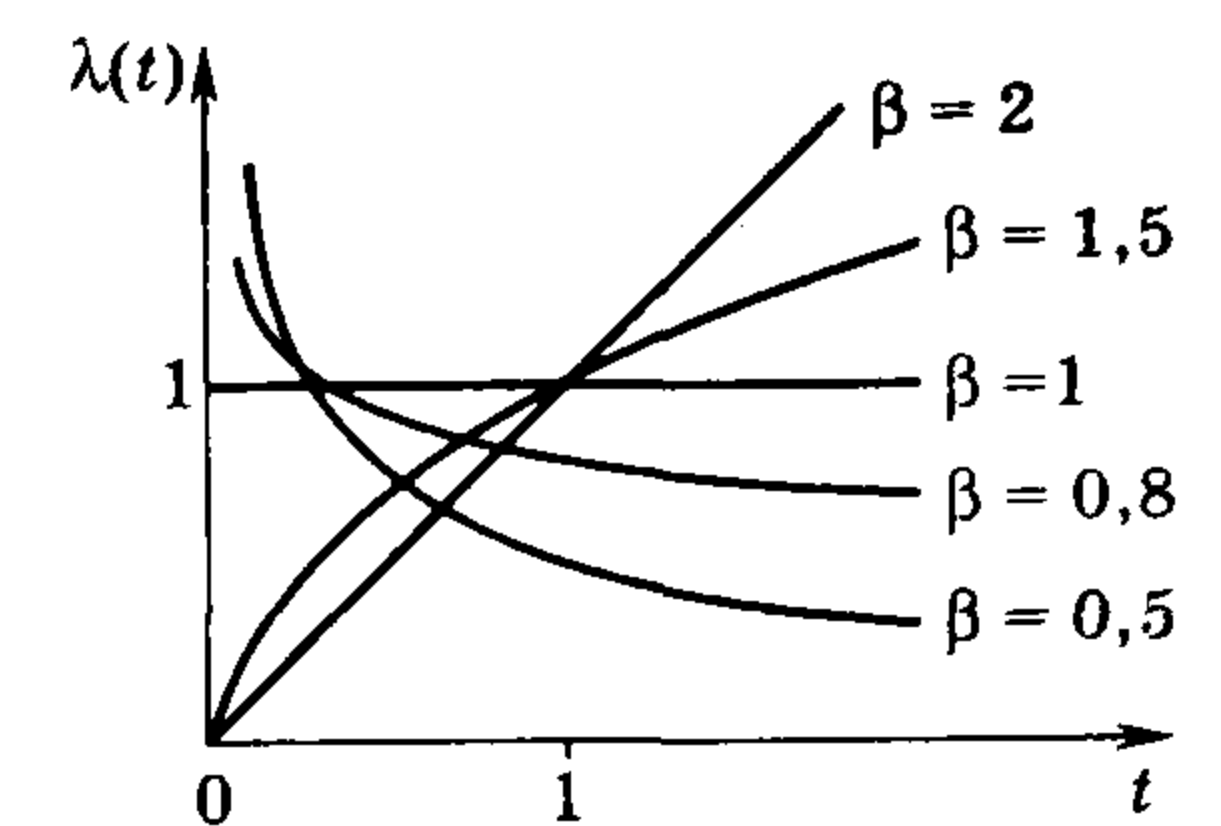


Рис. П 2.6

Распределение Вейбулла — Гнеденко обладает тем свойством, что его функция интенсивности

$$\lambda(t) = \lambda^\beta \beta t^{\beta-1} \quad (\text{П 2.52})$$

может быть как возрастающей (при $\beta > 1$), так и убывающей (при $\beta < 1$). Очевидно, что в случае $\beta = 1$ оно совпадает с экспоненциальным (рис. П 2.6).

Среднее этого распределения равно

$$E\{X\} = \frac{1}{\lambda} \Gamma\left(1 + \frac{1}{\beta}\right), \quad (\text{П 2.53})$$

а дисперсия

$$\sigma^2 = \frac{1}{\lambda^2} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2 \right], \quad (\text{П 2.54})$$

где Γ есть гамма-функция.

данных он на интуитивном уровне ввел двухпараметрическое распределение. Выдающийся русский математик Б. В. Гнеденко несколько позже дал строгое обоснование этого распределения. Он доказал ряд теорем, касающихся предельных распределений экстремальных случайных величин. Рассматриваемое распределение относится к одному из классов открытых им экстремальных распределений.

Решить эту систему алгебраических уравнений относительно каждого из p_k можно, используя правило Крамера:

$$p_k = \frac{D_k}{D}, \quad (\Pi 3.4)$$

где D — определитель системы алгебраических уравнений (П 3.13), а D_k — тот же определитель, в котором k -й столбец заменен на столбец свободных членов.

Обозначим через W множество всех состояний работоспособности системы. Тогда коэффициент готовности K , или вероятность того, что система в стационарном режиме находится в работоспособном состоянии, равняется

$$K = P_W = \sum_{k \in W} p_k = \frac{1}{D} \sum_{k \in W} D_k. \quad (\text{П } 3.5)$$

П 3.3. НАХОЖДЕНИЕ НЕСТАЦИОНАРНОГО КОЭФФИЦИЕНТА ГОТОВНОСТИ

Нестационарный КГ системы также находится на основании графа со всеми транзитивными состояниями (рис. П 3.1). Как и выше, $E(k)$ есть подмножество состояний, из которых возможен непосредственный переход в состояние k , а $e(k)$ есть подмножество состояний, в одно из которых процесс может перейти за бесконечно малый промежуток времени из состояния k . В отличие от предыдущего случая поток начинается в момент $t = 0$, для которого должно быть выбрано начальное состояние.

На основании графа переходов запишем формулу полной вероятности для состояния k :

$$P_k(t + \Delta) = P_k(t) \cdot \left[1 - \Delta \sum_{i \in \mathbf{e}(k)} \lambda_{ki} \right] + \Delta \sum_{i \in \mathbf{E}(k)} \lambda_{jk} P_j(t). \quad (\Pi 3.6)$$

Действительно, для того чтобы система в момент времени $t + \Delta$ оказалась в состоянии k , необходимо, чтобы она находилась в этом состоянии в момент t и не ушла бы из этого состояния за время Δ , либо чтобы она была в одном из смежных транзитивных состояний в момент t и перешла бы оттуда в состояние k за время Δ . Преобразуем (П 3.6) к виду

$$\frac{P_k(t + \Delta) - P_k(t)}{\Delta} = -P_k(t) \sum_{i \in e(k)} \lambda_{ki} + \sum_{j \in E(k)} \lambda_{jk} P_j(t). \quad (\Pi 3.7)$$

Предельный переход при $\Delta \rightarrow 0$ приводит к дифференциальному уравнению для состояния k :

$$\frac{dP_k(t)}{dt} = -P_k(t) \sum_{i \in e(k)} \lambda_{ki} + \sum_{j \in E(k)} \lambda_{jk} P_j(t). \quad (\Pi 3.8)$$

Если система имеет n^* различных транзитивных (непоглощающих) состояний в фазовом пространстве переходов, то для решения выбираются любые $(n^* - 1)$ уравнений для этих состояний и к ним добавляется условие нормировки

$$\sum_{k=1}^n p_k(t) = 1. \quad (\Pi 3.9)$$

При анализе динамики поведения нестационарного КГ выбирается начальное условие $P_k(0)=1$ для того состояния k , которое предполагается исходным при анализе. Иначе говоря, k – это индекс состояния, в котором система находится в момент $t = 0$. Обычно таким состояниям выбирается состояние полной работоспособности всех элементов системы, т. е. $P_0(0) = 1$. Преобразование Лапласа для уравнения (П 3.8) с учетом начального условия примет вид:

$$s\varphi_k(s) - P_k(0) = -\varphi_k(s) \sum_{i \in e(k)} \lambda_{ki} + \sum_{j \in E(k)} \lambda_{jk} \varphi_j(s), \quad (\Pi 3.10)$$

где как всегда, $\phi_k(s)$ есть преобразование Лапласа функции $P_k(t)$. Рассматривая систему алгебраических уравнений в преобразованиях Лапласа, берем те же $(n^* - 1)$ уравнений и к ним добавляем уравнение

$$s \sum_{k=1}^n \phi_k(s) = 1. \quad (\Pi 3.11)$$

В канонической форме система уравнений (П 3.10), (П 3.11) имеет вид

$$\begin{aligned} & -\left(s + \sum_{\forall j} \lambda_{1j}\right) \varphi_1(s) + \lambda_{21} \varphi_2(s) + \dots + \lambda_{n1} \varphi_n(s) = 1, \\ & \lambda_{12} \varphi_1(s) - \left(s + \sum_{\forall j} \lambda_{2j}\right) \varphi_2(s) + \dots + \lambda_{n2} \varphi_n(s) = 0, \\ & \dots\dots\dots \quad (\text{П 3.12}) \\ & \lambda_{1; n-1} \varphi_1(s) + \lambda_{2; n-1} \varphi_2(s) + \dots - \left(s + \sum_{\forall j} \lambda_{n-1; j}\right) \varphi_{n-1}(s) + \\ & \quad + \lambda_{n; n-1} \varphi_n(s) = 0, \\ & s \varphi_1(s) + s \varphi_2(s) + \dots + s \varphi_n(s) = 1. \end{aligned}$$

Решение системы (П 3.12) ищется с использованием правила Крамера

$$\varphi_k(s) = \frac{D_k(s)}{D(s)}, \quad (\Pi 3.13)$$

где $D(s)$ — определитель системы алгебраических уравнений (П 3.13), а $D_k(s)$ — тот же определитель, в котором k -й столбец заменен на столбец свободных членов.

Искомое преобразование Лапласа для вероятности того, что система, находясь в подмножестве состояний работоспособности W , не перейдет в какое-либо состояние отказа, равняется

$$\varphi_{\mathbf{W}}(s) = \sum_{k \in \mathbf{W}} \varphi_k(s) = \frac{1}{D(s)} \sum_{k \in \mathbf{W}} D_k(s). \quad (\Pi 3.14)$$

Обращение полученного преобразования Лапласа можно осуществить с помощью следующей простой процедуры.

■ Записываем $\varphi_{\mathbf{w}}(s)$ в виде

$$\varphi_w(s) = \frac{A_0 + A_1 s + \dots + A_n s^n}{B_0 + B_1 s + \dots + B_{n+1} s^{n+1}}, \quad (\Pi 3.15)$$

где A_k и B_k — известные коэффициенты, которые находятся после раскрытия определителей и приведения подобных членов.

■ Определяем корни полинома, стоящего в знаменателе В (П 3.15):

$$B_0 + B_1 s + \dots + B_{n+1} s^{n+1} = 0. \quad (\text{п } 3.16)$$

Пусть эти корни равны s_1, s_2, \dots, s_n . Это значит, что

$$B_0 + B_1 s + \dots + B_{n+1} s^{n+1} = \prod_{j=1}^{n+1} (s - s_j). \quad (\text{п } 3.17)$$

■ **Записываем $\phi_w(s)$ условно в виде суммы простых дробей:**

$$\varphi_W(s) = \frac{\beta_1}{s-s_1} + \frac{\beta_2}{s-s_2} + \dots + \frac{\beta_{n+1}}{s-s_{n+1}}, \quad (\Pi 3.18)$$

где β_k — искомые коэффициенты.

■ **Записываем $\varphi_w(s)$ в приведенной форме:**

$$\varphi_{\mathbf{w}}(s) = \frac{\sum_{1 \leq j \leq n} \beta_j \prod_{i \neq j} (s - s_i)}{\prod_{1 \leq i \leq n+1} (s - s_i)} \quad (\Pi 3.19)$$

и далее, раскрывая произведения и приводя подобные члены, получаем

$$\varphi_W(s) = \frac{\gamma_0 + \gamma_1 s + \gamma_2 s^2 + \dots + \gamma_n s^n}{(s - s_1)(s - s_2) \dots (s - s_{n+1})}, \quad (\text{П 3.20})$$

где γ_k выражены через различные β_j и s_j .

■ Полиномы вида (П 3.20) и (П 3.15) равны тогда и только тогда, когда $A_k = \gamma_k$. Из этих уравнений и *находятся* искомые коэффициенты β_k .

■ После нахождения β_k к преобразованиям (П 3.18) *применяется* обратное преобразование Лапласа (табл. 4.1, с. 67).

Замечание. Если корни оказываются кратными, т. е., например, значение s_j встречается k_j раз, то необходимо использовать обращение для преобразования Лапласа $\beta/(s - s_j)^{k_j}$.

П 3.4. ОЦЕНКА ВЕРОЯТНОСТИ БЕЗОТКАЗНОЙ РАБОТЫ

Рассмотрим систему в момент $t = 0$, когда все ее элементы изначально работоспособны. Пусть m состояний являются состояниями работоспособности, и их индексы принадлежат некоторому подмножеству W . Для вычисления ВБР составляется система дифференциальных уравнений, аналогичная рассмотренной выше. Отличие состоит в том, что все состояния отказа системы полагаются поглощающими. Начальное условие выбирается в виде $P_0(0) = 1$, т. е. предполагается, что начальное состояние «0» является состоянием полной работоспособности. Если система имеет m состояний работоспособности, то можно записать систему из m уравнений в преобразованиях Лапласа:

$$\begin{aligned} & -\left(s + \sum_{\forall j} \lambda_{1j}\right) \varphi_1(s) + \lambda_{21} \varphi_2(s) + \dots + \lambda_{m1} \varphi_m(s) = 1, \\ & \lambda_{12} \varphi_1(s) - \left(s + \sum_{\forall j} \lambda_{2j}\right) \varphi_2(s) + \dots + \lambda_{m2} \varphi_m(s) = 0, \\ & \dots\dots\dots \\ & \lambda_{1m} \varphi_1(s) + \lambda_{2m} \varphi_2(s) + \dots - \left(s + \sum_{\forall i} \lambda_{mi}\right) \varphi_m(s) = 0. \end{aligned} \quad (\Pi \text{ 3.21})$$

Заметим, что при решении этой задачи необходимость в использовании нормировочного уравнения отпадает. Интересующее нас преобразование Лапласа имеет вид

$$\varphi_W(s) = \sum_{j \in W} \varphi_j(s). \quad (\text{П } 3.22)$$

Формально нахождение вероятностей $p_k(t)$ ничем не отличается от схемы, предложенной для нахождения нестационарного коэффициента готовности. В этом случае условная вероятность безотказной работы системы при условии, что в начальный момент она находится в одном из состояний работоспособности, равна

$$P_W(t) = \sum_{j \in W} p_j(t), \quad (\text{П } 3.23)$$

где $p_j(t)$ находятся в результате решения системы (П 3.21).

П. 3.5. ОЦЕНКА СТАЦИОНАРНОЙ ВЕРОЯТНОСТИ БЕЗОТКАЗНОЙ РАБОТЫ

Рассмотрим ВБР системы, начиная с некоторого момента времени в стационарном режиме, при условии, что в этот начальный момент система находится в одном из состояний работоспособности. Находя выражение для коэффициента готовности системы, мы попутно определили все стационарные вероятности для всех n^* возможных транзитивных состояний системы. Пусть m из этих состояний являются состояниями работоспособности и их индексы принадлежат некоторому подмножеству W .

Вероятность того, что система в произвольно выбранный момент в стационарном режиме находится в одном из состояний работоспособности, равна

$$P_W = \sum_{j \in W} p_j. \quad (\text{П } 3.24)$$

Вероятность того, что система при этом условии находится именно в работоспособном состоянии k , $k \in W$, равна

$$p_k^* = \frac{p_k}{P_W}. \quad (\text{П } 3.25)$$

Для нахождения вероятности безотказной работы составляет система дифференциальных уравнений, аналогичная той, которая рассматривалась в предыдущем случае. Опять все состоя-

ния отказа делаются поглощающими. Однако в качестве начальных условий выбираются величины p_k^* . Если система имеет m состояний работоспособности, то можно записать систему из m дифференциальных уравнений в преобразованиях Лапласа:

$$\begin{aligned} -\left(s + \sum_{j \in W} \lambda_{1j}\right) \varphi_1(s) + \lambda_{21} \varphi_2(s) + \dots + \lambda_{m1} \varphi_m(s) &= p_1^*, \\ \lambda_{12} \varphi_1(s) - \left(s + \sum_{j \in W} \lambda_{2j}\right) \varphi_2(s) + \dots + \lambda_{m2} \varphi_m(s) &= -p_2^*, \\ &\dots\dots\dots \\ \lambda_{1m} \varphi_1(s) + \lambda_{2m} \varphi_2(s) + \dots - \left(s + \sum_{j \in W} \lambda_{mj}\right) \varphi_m(s) &= -p_m^*. \end{aligned} \quad (\text{П } 3.26)$$

И в этом случае необходимость в использовании нормировочного уравнения отпадает. Интересующее нас преобразование Лапласа имеет вид

$$\varphi_W(s) = \sum_{j \in W} \varphi_j(s). \quad (\text{П } 3.27)$$

Формально нахождение вероятностей $p_k(t)$ ничем не отличается от приведенного выше. В этом случае условная вероятность безотказной работы системы при условии, что в начальный момент она находится в одном из состояний работоспособности, равна

$$P_W(t) = \sum_{j \in W} p_j(t), \quad (\text{П } 3.28)$$

где $p_j(t)$ определяется в результате решения системы (П 3.27).

П 3.6. НАХОЖДЕНИЕ СРЕДНЕГО ВРЕМЕНИ РАБОТЫ ДО ОЧЕРЕДНОГО ОТКАЗА И МЕЖДУ ОТКАЗАМИ

Известно, что если $\varphi(s)$ есть преобразование Лапласа для ВБР $P(t)$, то среднее время безотказной работы можно записать в виде

$$T = \int_0^{\infty} P(t) dt = \left[\int_0^{\infty} P(t) e^{-st} dt \right]_{s=0} = \varphi(s) \Big|_{s=0}. \quad (\text{П } 3.29)$$

Если нас интересует значение среднего времени работы до отказа, то следует взять решение (П 3.22) и сделать подстановку $s = 0$. Если же мы ищем среднее время между отказами для стационарного режима, то нужно взять решение (П 3.27) и в нем сделать подстановку $s = 0$.

П 3.7. НАХОЖДЕНИЕ СРЕДНЕГО ВРЕМЕНИ ВОССТАНОВЛЕНИЯ

Нахождение этого показателя аналогично нахождению среднего времени безотказной работы, но при переопределении подмножества транзитивных и поглощающих состояний. В этом случае в исходном графе переходов все состояния работоспособности нужно сделать поглощающими и рассматривать процесс блуждания внутри подмножества состояний отказа w до попадания его в одно из состояний подмножества W . Если общее число состояний системы n , а число состояний работоспособности m , то для оставшихся $k = (n - m)$ состояний составляются k уравнений

$$\begin{aligned} & -\left(s + \sum_{j=1}^n \lambda_{1j}\right) \varphi_1(s) + \lambda_{21} \varphi_2(s) + \dots + \lambda_{k1} \varphi_k(s) = q_1^*, \\ & \lambda_{12} \varphi_1(s) - \left(s + \sum_{j=2}^n \lambda_{2j}\right) \varphi_2(s) + \dots + \lambda_{k2} \varphi_k(s) = q_2^*, \\ & \dots\dots\dots \\ & \lambda_{1k} \varphi_1(s) + \lambda_{2k} \varphi_2(s) + \dots - \left(s + \sum_{j=k}^n \lambda_{kj}\right) \varphi_k(s) = q_k^*, \end{aligned} \quad (\Pi 3.30)$$

где каждая из величин q_k^* , $k \in w$, определяется как

$$q_k^* = \frac{p_j}{\sum_{j \in W} p_j}. \quad (\Pi 3.31)$$

После решения этой системы среднее время восстановления находится по формуле

$$\tau = \sum_{j \in W} \varphi_j(0). \quad (\Pi 3.32)$$

Приложение 4

Процесс гибели и размножения

Процесс *гибели и размножения* широко используется в теории надежности и теории массового обслуживания. Он является частным случаем дискретного марковского процесса, а именно однородным марковским процессом с конечным или счетным числом состояний, которые мы будем обозначать натуральными числами $0, 1, 2, \dots$. Если в момент t процесс находится в состоянии k , то за бесконечно малый промежуток времени Δ он с вероятностью $\lambda_k \Delta + o(\Delta)$ перейдет в состояние $(k + 1)$, с вероятностью $\mu_k \Delta + o(\Delta)$ перейдет в состояние $(k - 1)$, а с вероятностью $1 - (\lambda_k + \mu_k) \Delta + o(\Delta)$ останется в состоянии k . Заметим, что состояния 0 и n являются *отражающими*, т. е. $\lambda_0 = 0$ и $\mu_n = 0$. Соответствующий граф переходов изображен на рис. П 4.1.

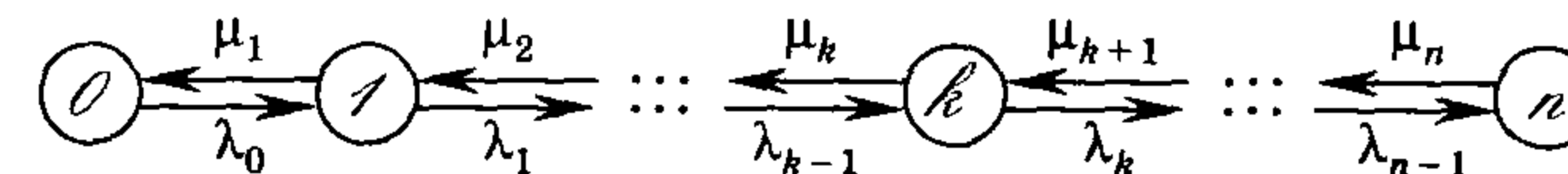


Рис. П 4.1

Для состояния k можно записать следующее условие баланса:

$$P_k(t + \Delta) = P_{k-1}(t)[\Delta + o(\Delta)] + \\ + P_k(t)[1 - (\lambda_k + \mu_k)\Delta + o(\Delta)] + P_{k+1}(t)[\mu_k\Delta + o(\Delta)]. \quad (\Pi 4.1)$$

Перегруппировав члены уравнения (П 4.1), получим

$$\frac{P_k(t + \Delta) - P_k(t)}{\Delta} = \lambda_{k-1}P_{k-1}(t) - P_k(t)(\lambda_k + \mu_k) + \mu_k P_{k+1}(t) + o(\Delta), \quad (\Pi 4.2)$$

что при $\Delta \rightarrow 0$ приводит к дифференциальному уравнению

$$\frac{d}{dt} P_k(t) = \lambda_{k-1} P_{k-1}(t) - P_k(t) (\lambda_k + \mu_k) + \mu_k P_{k+1}(t). \quad (14.3)$$

Аналогичным образом выводится и уравнение для состояния 0 :

$$\frac{d}{dt}P_0(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t). \quad (\Pi 4.4)$$

Если процесс имеет конечное число состояний n , то последнее уравнение имеет вид

$$\frac{d}{dt} P_n(t) = \lambda_{n-1} P_{n-1}(t) - \mu_n P_n(t). \quad (\text{П 4.5})$$

Если положить $\mu_0 = \lambda_{-1} = 0$ (а в случае конечного числа состояний еще и $\lambda_n = \mu_{n+1} = 0$), то можно обойтись выражением (П 4.3). В итоге можно записать следующую систему дифференциальных уравнений для конечного числа состояний:

$$\begin{aligned} \frac{d}{dt} P_0(t) &= -\lambda_0 P_0(t) + \mu_1 P_1(t), \\ \frac{d}{dt} P_1(t) &= \lambda_0 P_0(t) - P_1(t)(\lambda_1 + \mu_1) + \mu_2 P_2(t), \\ &\dots\dots\dots \\ \frac{d}{dt} P_k(t) &= \lambda_{k-1} P_{k-1}(t) - P_k(t)(\lambda_k + \mu_k) + \mu_k P_{k+1}(t), \\ &\dots\dots\dots \\ \frac{d}{dt} P_n(t) &= \lambda_{n-1} P_{n-1}(t) - \mu_n P_n(t). \end{aligned} \quad (\text{П 4.6})$$

В случае счетного числа состояний решение производится для бесконечного числа уравнений типа (П 4.6).

При решении системы дифференциальных уравнений необходимо задать также и начальные условия, которые обычно представляются в виде $P_0(0) = 1$.

Физическая интерпретация уравнений баланса была приведена в Приложении 3. Повторим вкратце эти рассуждения. Представим каждое состояние системы k неким резервуаром с объемом жидкости $p_k(t)$ в момент t . Единица объема жидкости из резервуара k вытекает в резервуар $(k-1)$ с интенсивностью μ_k и в резервуар $(k+1)$ с интенсивностью λ_k . В то же время сам резервуар k по такому же правилу пополняется из соседних резервуаров $(k-1)$ и $(k+1)$. Интуитивно понятно, что при таком перетекании жидкости в конце концов наступит стационарное равновесие, когда объем жидкости в каждом резервуаре уже не будет изменяться. Это означает, что все производные $\frac{d}{dt} P_k(t) \rightarrow 0$ при $t \rightarrow \infty$.

Обычно на практике интересуются поведением процесса на промежутках времени, далеко удаленных от начального момента, т. е. именно стационарными характеристиками. Тогда система

дифференциальных уравнений вырождается в систему алгебраических уравнений:

$$\begin{aligned} 0 &= -\lambda_0 p_0 + \mu_1 p_1, \\ 0 &= \lambda_0 p_0 - (\lambda_1 + \mu_1) p_1 + \mu_2 p_2, \\ &\dots\dots\dots \\ 0 &= \lambda_{k-1} p_{k-1} - (\lambda_k + \mu_k) p_k + \mu_k p_{k+1}, \\ &\dots\dots\dots \\ 0 &= \lambda_{n-1} p_{n-1} - \mu_n p_n. \end{aligned} \quad (\text{П 4.7})$$

Дополнительно к этому используем условие полной группы событий

$$\sum_{k=0}^n p_k = 1. \quad (\text{П 4.8})$$

Уравнения баланса в схеме *гибели* и *размножения* удобнее записывать не для состояний, а для *разрезов* графа переходов (рис. П 4.2).

Действительно, состояние баланса характеризуется тем, что в стационарном режиме в каждом разрезе интенсивность потока слева направо равна интенсивности потока справа налево. Поэтому можно записать уравнения баланса для разрезов 1, 2 и т. д. в виде

$$\begin{aligned} \lambda_0 p_0 &= \mu_1 p_1, \\ \lambda_1 p_1 &= \mu_2 p_2, \\ &\dots\dots\dots \\ \lambda_{k-1} p_{k-1} &= \mu_k p_k, \\ &\dots\dots\dots \\ \lambda_{n-1} p_{n-1} &= \mu_n p_n. \end{aligned} \quad (\text{П 4.9})$$

Отсюда, обозначив $\rho_k = \lambda_k / \mu_k$, находим последовательно:

$$\begin{aligned} p_1 &= \rho_1 p_0, \\ p_2 &= \rho_2 p_1 = \rho_1 \rho_2 p_0, \\ &\dots\dots\dots \\ p_k &= p_0 \prod_{j=1}^k \rho_j, \\ &\dots\dots\dots \\ p_n &= p_0 \prod_{j=1}^n \rho_j. \end{aligned} \quad (\text{П 4.10})$$

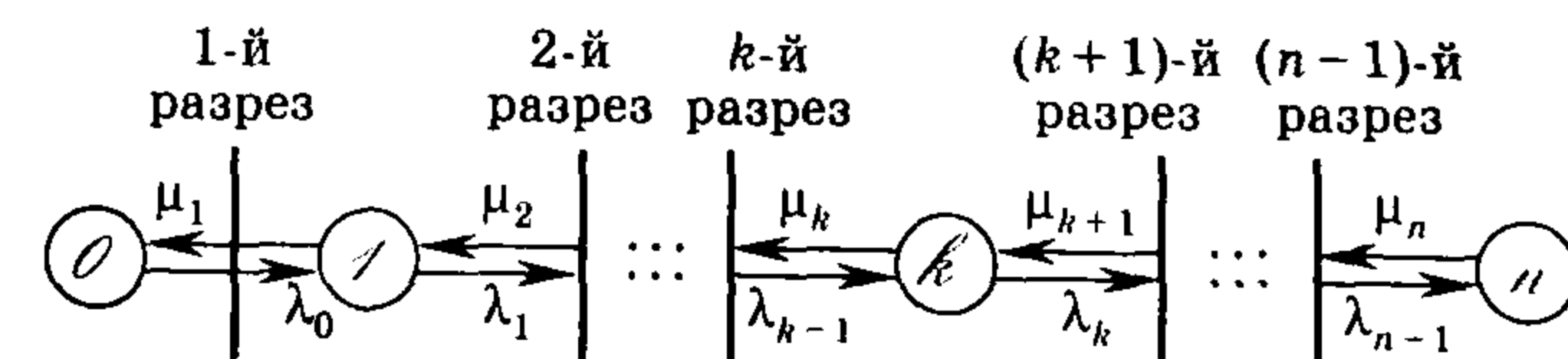


Рис. П 4.2

Используя условие (П 4.8), можно найти

$$p_0 = \left(1 + \sum_{k=1}^n \prod_{j=1}^k \rho_j \right)^{-1}. \quad (\text{П 4.11})$$

В результате получаем при любом n выражение стационарной вероятности p_k :

$$p_k = \frac{\prod_{j=1}^k \rho_j}{1 + \sum_{k=1}^n \prod_{j=1}^k \rho_j}. \quad (\text{П 4.12})$$

Если рассматривается схема *гибели и размножения* для счетного числа состояний, формула (П4.12) естественным образом трансформируется в формулу

$$p_k = \frac{\prod_{j=1}^k \rho_j}{1 + \sum_{k=1}^{\infty} \prod_{j=1}^k \rho_j}. \quad (\text{П 4.13})$$

В этом случае стационарное решение существует, если последовательность

$$\rho^{(k)} = \prod_{j=1}^k \rho_j \quad (\text{П 4.14})$$

является монотонно убывающей с ростом k и в пределе стремится к нулю.

Приложение 5

Многоканальная система массового обслуживания

П 5.1. РАСПРЕДЕЛЕНИЕ ВЕРОЯТНОСТЕЙ ЗАНЯТЫХ КАНАЛОВ

В Приложении 4 была рассмотрена схема *гибели и размножения* в общем виде. Здесь нам понадобится одна из конкретных схем такого типа: многоканальная система массового обслуживания с постоянным входным потоком. Такая система описывает работу ремонтной базы в случае, когда из-за большого общего числа объектов можно полагать, что поток поступающих на ремонтную базу отказавших элементов пуассоновский с постоянной интенсивностью.

Пусть известны следующие параметры:

- ✓ Λ — входной поток отказавших элементов от всех объектов системы;
- ✓ μ — интенсивность ремонта на каждом ремонтном органе;
- ✓ d — число ремонтных органов на ремонтной базе.

Как известно, для того чтобы на входе ремонтной базы не накапливалась бы неограниченная очередь, должно выполняться условие $\Lambda < d\mu$.

Рассмотрим соответствующую марковскую модель (рис. П 5.1). Обозначим через k состояния системы, когда в ней находится k элементов (в ремонте или в очереди на ремонт). Поток поступающих в систему отказов можно считать практически неограниченным. При наличии в системе массового обслуживания d или менее отказавших элементов ремонтируются все они одновременно. Если число отказавших элементов на ремонтной базе равно k , $k > d$, то $(k - d)$ из них образуют общую очередь. Система алгебраических

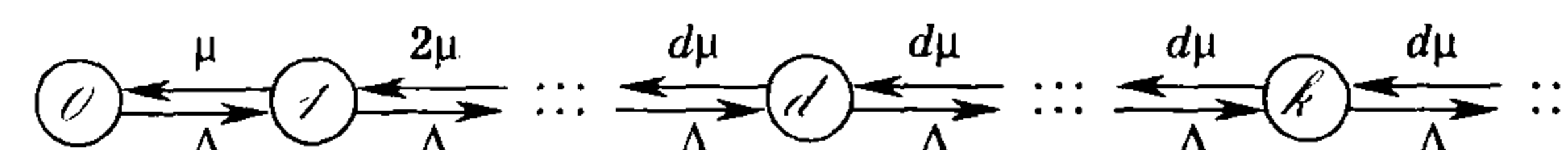


Рис. П 5.1

уравнений, описывающих стационарные состояния этой системы массового обслуживания, имеет вид

$$\begin{aligned} \Lambda p_0 &= \mu p_1, \\ \Lambda p_1 &= 2\mu p_2, \\ &\dots\dots\dots \\ \Lambda p_{d-1} &= d\mu p_d, \\ \Lambda p_d &= d\mu p_{d+1}, \\ &\dots\dots\dots \\ \Lambda p_{k-1} &= d\mu p_k, \\ \Lambda p_k &= d\mu p_{k+1}, \\ &\dots\dots\dots \end{aligned} \quad (\text{П } 5.1)$$

Обозначим $\rho = \Lambda/\mu$. Тогда получим решение в виде

$$\begin{aligned} p_1 &= \rho p_0, \\ p_2 &= \frac{\rho}{2} p_1 = \frac{\rho^2}{2} p_0, \\ p_3 &= \frac{\rho}{3} p_2 = \frac{\rho^3}{2 \cdot 3} p_0, \\ &\dots\dots\dots \\ p_d &= \frac{\rho}{d} p_{d-1} = \frac{\rho^d}{d!} p_0, \\ p_{d+1} &= \frac{\rho}{d} p_d = \frac{\rho^{d+1}}{d \cdot d!} p_0, \\ &\dots\dots\dots \\ p_k &= \frac{\rho}{d} p_{k-1} = \frac{\rho^k}{d^{k-1} \cdot d!} p_0, \\ p_{k+1} &= \frac{\rho}{d} p_k = \frac{\rho^{k+1}}{d^{k-d} \cdot d!} p_0, \\ &\dots\dots\dots \end{aligned} \quad (\text{П } 5.2)$$

Из условия полной группы событий

$$\sum_{k=0}^{\infty} p_k = 1$$

легко находится вероятность p_0 :

$$p_0 = \frac{1}{1 + \sum_{j=1}^d \frac{\rho^j}{j!} + \sum_{j=d+1}^{\infty} \frac{\rho^j}{d^{j-d} \cdot d!}}. \quad (\text{П } 5.3)$$

В предположении, что $\rho < d$, вторую сумму, стоящую в знаменателе, можно записать в виде

$$\sum_{j=d+1}^{\infty} \frac{\rho^j}{d^{j-d} \cdot d!} = \frac{d^d}{d!} \sum_{j=d+1}^{\infty} \frac{\rho^j}{d^j} = \frac{\rho^{d+1}}{d!(d-\rho)}. \quad (\text{П } 5.4)$$

В результате, учитывая, что $\frac{\rho^0}{0!} = 1$, можно (П 5.3) переписать в виде

$$p_0 = \frac{1}{\sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d-\rho)}}. \quad (\text{П } 5.5)$$

На основании (П 5.5) из (П 5.2) путем подстановки легко находятся все остальные вероятности p_k ; для $j \leq d$ и для $j > d$ соответственно:

$$p_k = \frac{\rho^k}{d! \left[\sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d-\rho)} \right]}, \quad (\text{П } 5.6)$$

$$p_k = \frac{\rho^k}{d^{k-d} d! \left[1 + \sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d-\rho)} \right]}. \quad (\text{П } 5.7)$$

П 5.2. ОПРЕДЕЛЕНИЕ СРЕДНЕГО ВРЕМЕНИ ОЖИДАНИЯ

В Приложении 4 была рассмотрена схема гибели и размножения для d -канальной системы массового обслуживания и получена вероятность того, что в такой системе будет находиться k заявок на обслуживании:

$$p_k = \frac{\rho^k}{k!} p_0, \quad k \leq d, \quad (\text{П } 5.8)$$

$$p_k = \left(\frac{\rho}{d}\right)^{k-d}, \quad p_d = \frac{\rho^d}{d^{d-1} \cdot d!} p_0, \quad k > d, \quad (\text{П } 5.9)$$

где

$$p_0 = \frac{1}{\sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d-\rho)}} \quad (\text{П } 5.10)$$

и $\rho = \Lambda/\mu$, причем Λ — интенсивность входящего потока, μ — интенсивность обслуживания.

Для дальнейшего понадобится еще и вероятность π того, что все обслуживающие каналы заняты:

$$\pi = \sum_{k=d}^{\infty} p_k = \frac{d^d}{d!} p_0 \sum_{k=d}^{\infty} \left(\frac{\rho}{d}\right)^k = \frac{\rho^d p_0}{(d-1)!(d-\rho)}. \quad (\text{П } 5.11)$$

Обозначим через γ случайную длительность ожидания обслуживания. Тогда $P\{\gamma > t\}$ есть вероятность того, что длительность ожидания превзойдет t . Обозначим через $P_k\{\gamma > t\}$ вероятность неравенства $\gamma > t$ при условии, что в момент поступления очередной

заявки в системе уже находилось k заявок. По формуле полной вероятности справедливо равенство

$$P\{\gamma > t\} = \sum_{k=0}^{\infty} p_k P_k\{\gamma > t\}. \quad (\text{П } 5.12)$$

При любых $k < d$ и любых $t > 0$ выполняется равенство $P_k\{\gamma > t\} = 0$, т. е. можно записать

$$P\{\gamma > t\} = \sum_{k=d}^{\infty} p_k P_k\{\gamma > t\}. \quad (\text{П } 5.13)$$

Если в момент поступления очередной заявки в очереди уже находилось $k = d$ заявок, то при дисциплине FIFO* эта заявка будет ожидать, пока не будет обслужено $(k - d + 1)$ заявок. Пусть $q_s(t)$ обозначает вероятность того, что за время t с момента поступления заявки обслужено s заявок. При $k > d$ имеет место неравенство

$$P_k\{\gamma > t\} = \sum_{s=0}^{k-d} q_s(t). \quad (\text{П } 5.14)$$

На интервале времени со всеми занятыми каналами поток обслуженных заявок образует пуассоновский поток с интенсивностью μd . Это означает, что вероятность освобождения s каналов равна

$$q_s(t) = \frac{(d\mu t)^s}{s!} \exp(-\mu d t). \quad (\text{П } 5.15)$$

Подставляя (П 5.15) в (П 5.14), находим

$$P_k\{\gamma > t\} = \sum_{s=0}^{k-d} \frac{(d\mu t)^s}{s!} \exp(-\mu d t) \quad (\text{П } 5.16)$$

и, следовательно,

$$P\{\gamma > t\} = \sum_{k=d}^{\infty} p_k \sum_{s=0}^{k-d} \frac{(d\mu t)^s}{s!} \exp(-\mu d t). \quad (\text{П } 5.17)$$

Подставляя p_k из (П 5.9) в (П 5.17), получаем

$$P\{\gamma > t\} = p_d \exp(-\mu d t) \sum_{k=d}^{\infty} \left(\frac{\rho}{d}\right)^{k-d} \sum_{s=0}^{k-d} \frac{(d\mu t)^s}{s!}. \quad (\text{П } 5.18)$$

После несложных преобразований имеем

$$\begin{aligned} P\{\gamma > t\} &= p_d \exp(-\mu d t) \sum_{s=0}^{\infty} \frac{(d\mu t)^s}{s!} \sum_{k=d+s}^{\infty} \left(\frac{\rho}{d}\right)^{k-d} = \\ &= p_d \exp(-\mu d t) \sum_{s=0}^{\infty} \frac{(d\mu t)^s}{d^s s!} \sum_{k=d+s}^{\infty} \left(\frac{\rho}{d}\right)^{k-d-s} = \\ &= \frac{p_d}{1 - \frac{\rho}{d}} \exp(-\mu d t) \sum_{s=0}^{\infty} \frac{(\Lambda t)^s}{s!} = \frac{p_d}{1 - \frac{\rho}{d}} \exp[-(\mu d - \Lambda t)]. \end{aligned} \quad (\text{П } 5.19)$$

См. сноску к с. 194. — Ред.

Вспомним теперь, что

$$P_d = \frac{\rho^d}{d!} p_0 = \pi \left(1 - \frac{\rho}{d}\right). \quad (\text{П } 5.20)$$

Таким образом,

$$P\{\gamma > t\} = \pi \exp[-(d\mu - \Lambda t)]. \quad (\text{П } 5.21)$$

Заметим, что в точке $t = 0$ функция $P\{\gamma > t\}$ имеет разрыв непрерывности, равный вероятности π .

Теперь нетрудно найти и среднее время ожидания W :

$$W = \int_0^{\infty} P\{\gamma > t\} dt = \pi \int_0^{\infty} \exp[-(d\mu - \Lambda)t] dt = \frac{\pi}{\mu(d - \rho)}, \quad (\text{П } 5.22)$$

где π определено ранее в (П 5.11).

В развернутом виде формула для среднего времени ожидания в многоканальной системе массового обслуживания имеет вид

$$W = \frac{\pi}{\mu(d - \rho)} = \frac{\rho^d}{\mu(d - 1)!(d - \rho)^2} \left[\sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d - \rho)} \right]^{-1}. \quad (\text{П } 5.23)$$

Нам окажется полезным также иметь выражение для частного случая — одноканальной системы массового обслуживания. Сначала проведем предварительные выкладки, упростив выражения для p_0 и π при $d = 1$:

$$p_0 = \frac{1}{\left[1 + \sum_{j=0}^d \frac{\rho^j}{j!} + \frac{\rho^{d+1}}{d!(d - \rho)}\right]} = \frac{1}{1 + \rho + \frac{\rho^2}{1 - \rho}} = 1 - \rho, \quad (\text{П } 5.24)$$

$$\pi = \frac{\rho^d p_0}{(d - 1)!(d - \rho)} = \frac{\rho(1 - \rho)}{1 - \rho} = \rho. \quad (\text{П } 5.25)$$

В результате получаем для одноканальной системы выражение среднего времени ожидания в виде

$$W = \frac{\pi}{\mu(d - \rho)} = \frac{\rho}{\mu(1 - \rho)} = \frac{\Lambda}{\mu(\mu - \Lambda)}. \quad (\text{П } 5.26)$$

В тех случаях, когда предположение об экспоненциальности времени восстановления и о пуассоновском характере потока отказов оказываются на практике неприемлемыми, приходится прибегать к статистическому моделированию.

Оглавление

Предисловие	3
Глава 1. Показатели надежности	5
§ 1.1. Роль показателей надежности	5
§ 1.2. Невосстанавливаемые объекты	6
§ 1.3. Восстанавливаемые объекты	12
§ 1.4. Выбор показателей надежности	17
§ 1.5. Задание требований по надежности	18
Глава 2. Невосстанавливаемые системы	23
§ 2.1. Структурная функция	23
§ 2.2. Последовательная система	24
§ 2.3. Правило де Моргана	26
§ 2.4. Параллельная система — <i>горячий</i> резерв	27
§ 2.5. Система типа « <i>k</i> из <i>n</i> »	29
§ 2.6. Системы со смешанной структурой	30
§ 2.7. Ненагруженное резервирование — <i>холодный</i> резерв	33
§ 2.8. Облегченное резервирование — <i>теплый</i> резерв	35
Глава 3. Системы с сетевой структурой	36
§ 3.1. Мостиковая структура	36
§ 3.2. Метод вычисления функции связности	38
§ 3.3. Граничные оценки для мостиковой структуры	44
§ 3.4. Сетевые системы общего вида	51
§ 3.5. Декомпозиция сетевых структур	59
Глава 4. Марковские модели восстанавливаемых систем	64
§ 4.1. Пуассоновский поток	64
§ 4.2. Восстанавливаемый элемент	69
§ 4.3. Последовательная система	73
§ 4.4. Дублированная система	79
§ 4.5. Метод введения фиктивных состояний	88
Глава 5. Стареющие элементы и системы	90
§ 5.1. Монотонная функция интенсивности отказов	90
§ 5.2. <i>Стареющий</i> элемент	91
§ 5.3. Последовательная система из <i>стареющих</i> элементов	97
§ 5.4. Параллельные системы	101
§ 5.5. Границы для коэффициента интервальной готовности	105

Глава 6. Эвристические методы расчета надежности	107
§ 6.1. Одиночный элемент	107
§ 6.2. Предельная теорема о суперпозиции потоков	108
§ 6.3. Высоконадежная последовательная система	111
§ 6.4. Предельная теорема о разрежении потоков	113
§ 6.5. Высоконадежное <i>холодное</i> дублирование	116
§ 6.6. Высоконадежное <i>горячее</i> дублирование	119
§ 6.7. Элемент с периодическим контролем работоспособности	121
§ 6.8. Учет времени переключения при дублировании с <i>холодным</i> резервом	123
§ 6.9. Учет времени переключения при дублировании с <i>горячим</i> резервом	124
§ 6.10. Неполный контроль основного элемента дублированной системы с <i>горячим</i> резервом	125
§ 6.11. Неполный контроль резервного элемента дублированной системы с <i>горячим</i> резервом	127
§ 6.12. Учет ненадежности переключения	129
Глава 7. Оценка эффективности функционирования сложных систем	130
§ 7.1. Необходимость расширения понятия надежности	130
§ 7.2. Классы систем	131
§ 7.3. Системы кратковременного действия	132
§ 7.4. Системы с ветвящейся структурой	136
§ 7.5. Система с перекрывающимися зонами действия	142
Глава 8. Оптимальное резервирование	144
§ 8.1. Формулировка проблемы	144
§ 8.2. Метод множителей Лагранжа	147
§ 8.3. Метод покоординатного наискорейшего спуска	149
§ 8.4. Метод динамического программирования	158
§ 8.5. Оптимальное резервирование при нескольких ограничениях	167
Глава 9. Обеспечение систем типовыми элементами замены	169
§ 9.1. Общее описание системы	169
§ 9.2. Анализ индивидуального склада	172
§ 9.3. Анализ регионального склада	176
§ 9.4. Центральный склад	181
§ 9.5. Программа выбора оптимального числа запасных элементов для спутниковой телекоммуникационной сети «GlobalStar»	184
Глава 10. Территориально-распределенные системы технического обслуживания	190
§ 10.1. Описание анализируемой системы	190
§ 10.2. Зона с единственной передвижной ремонтной мастерской	191
§ 10.3. Зона обслуживания с несколькими передвижными ремонтными мастерскими	194
§ 10.4. Пример построения зон обслуживания передвижными ремонтными мастерскими	197

Приложение 1. Основные понятия и термины теории вероятностей	203
Приложение 2. Основные распределения, используемые в теории надежности	206
п 2.1. Распределение Бернулли	206
п 2.2. Геометрическое распределение	207
п 2.3. Экспоненциальное распределение	208
п 2.4. Гиперэкспоненциальное распределение	209
п 2.5. Биномиальное распределение	210
п 2.6. Распределение Пуассона	211
п 2.7. Распределение Эрланга	213
п 2.8. Нормальное распределение	214
п 2.9. Равномерное распределение	216
п 2.10. Вырожденное распределение	216
п 2.11. Отношения между распределениями вероятностей	216
п 2.12. Распределение Вейбулла — Гнеденко	218
Приложение 3. Сведения из теории марковских процессов	220
п 3.1. Общее описание	220
п 3.2. Нахождение стационарного коэффициента готовности	220
п 3.3. Нахождение нестационарного коэффициента готовности	222
п 3.4. Оценка вероятности безотказной работы	225
п 3.5. Оценка стационарной вероятности безотказной работы	226
п 3.6. Нахождение среднего времени работы до очередного отказа и между отказами	227
п 3.7. Нахождение среднего времени восстановления	228
Приложение 4. Процесс гибели и размножения	229
Приложение 5. Многоканальная система массового обслуживания	233
п 5.1. Распределение вероятностей занятых каналов	233
п 5.2. Определение среднего времени ожидания	235

Профессор **И. А. Ушаков** является признанным лидером в области теории надежности и ее промышленных приложений. Он заведовал кафедрами в МФТИ и МЭИ, где читал курсы по исследованию операций, теории надежности и дискретной оптимизации, а также работал во многих научно-исследовательских промышленных организациях и в АН СССР. На базе своего опыта работы в промышленности ученый опубликовал фундаментальный «Справочник по надежности» (1966 г.), многократно переизданный в нашей стране, США, ГДР, ФРГ и Чехословакии.

В дальнейшем профессор **И. А. Ушаков** преподавал в университетах США: Джорджа Вашингтона, Джорджа Мейсона и Калифорнийском университете Сан-Диего, а также выступал в качестве эксперта по надежности в различных американских компаниях.

-
- Настоящее учебное пособие предназначено как для инженерных вузов, так и для математических факультетов университетов. Будущие инженеры получают благодаря ему необходимые теоретические знания, а будущие математики-прикладники — представление о сути инженерных задач и методах их формализации.
 - Для работы с книгой не понадобится специальных знаний, выходящих за рамки стандартных математических курсов. Необходимые дополнительные сведения имеются в приложениях. Приводится много примеров, взятых из инженерной практики.

ISBN 978-5-358-01586-9



9 785358 015869