

И. Н. КОВАЛЕНКО
Г. К. МОСКАТОВ
Е. Ю. БАРЗИЛОВИЧ

**ПОЛУМАРКОВСКИЕ
МОДЕЛИ
В ЗАДАЧАХ
ПРОЕКТИРОВАНИЯ
СИСТЕМ УПРАВЛЕНИЯ
ЛЕТАТЕЛЬНЫМИ
АППАРАТАМИ**

ОСНОВЫ ПРОЕКТИРОВАНИЯ СИСТЕМ УПРАВЛЕНИЯ ЛЕТАТЕЛЬНЫМИ АППАРАТАМИ

СПРАВОЧНАЯ БИБЛИОТЕКА
ИНЖЕНЕРА-КОНСТРУКТОРА

Редколлегия

Заслуженный деятель науки и техники
РСФСР,

д-р техн. наук, проф. Б. А. Рябов;
заслуженный деятель науки и техники
РСФСР,

д-р техн. наук, проф. А. С. Шаталов;

д-р техн. наук, проф. В. А. Боднер;

д-р техн. наук, проф. Ю. И. Топчиев

И. Н. КОВАЛЕНКО,
Г. К. МОСКАТОВ,
Е. Ю. БАРЗИЛОВИЧ

ПОЛУМАРКОВСКИЕ
МОДЕЛИ
В ЗАДАЧАХ
ПРОЕКТИРОВАНИЯ
СИСТЕМ УПРАВЛЕНИЯ
ЛЕТАТЕЛЬНЫМИ
АППАРАТАМИ

Москва
«МАШИНОСТРОЕНИЕ»
1973

Коваленко И. Н., Москатов Г. К., Барзилович Е. Ю. Полумарковские модели в задачах проектирования систем управления летательными аппаратами. М., «Машиностроение», 1973, стр. 176.

В книге рассмотрены марковские процессы с конечным или счетным множеством состояний, полумарковские процессы вложенные цепи Маркова, процессы восстановления. Дана элементарная теория оптимальной остановки для случайной последовательности, служащая основанием некоторых методов оптимального управления.

Результаты теоретических исследований применены для решения отдельных задач, возникающих при проектировании и обслуживании систем управления летательными аппаратами.

Приведены методы расчета надежности адаптивных автопилотов, определения необходимого объема буферной памяти БЦВМ, оптимизации профилактического обслуживания САУ, управления посадкой групп самолетов, организации дозаправки самолетов в воздухе. Для иллюстрации методов расчета приведены соответствующие примеры.

Книга предназначена для инженеров, занимающихся проектированием систем автоматического управления летательными аппаратами.

Табл. 12. Ил. 40. Список лит. 71 назв.

К $\frac{331-215}{038(01)-73}$ 215-73

ПРЕДИСЛОВИЕ

Настоящая книга посвящена изложению математического подхода к решению ряда задач проектирования и организации обслуживания систем управления летательными аппаратами. Этот подход основан на теории массового обслуживания и теории полумарковских случайных процессов.

Работы А. А. Маркова об «испытаниях, связанных в цепь», легли в основу общей теории случайных процессов. Заслуживают также упоминания исследования «непрерывных вероятностей» французского математика Башелье, получившие прочную логическую основу лишь после создания теоретико-множественной системы построения основ теории вероятностей. Общая теория марковских процессов и их классификация были даны А. Н. Колмогоровым в 1930 г.

Теория массового обслуживания возникла в начале XX в. на базе задач телефонии. Позже выяснилось, что аналогичные или весьма близкие к ним задачи возникают на транспорте, при автоматизации технологических процессов, в теории исследования операций и в других областях науки и техники.

В монографии А. Я. Хинчина [46] были обобщены достигнутые к середине текущего столетия результаты в теории массового обслуживания (ТМО) и этой теории дана строгая научная основа. Заменяв разного рода эвристические рассуждения, применявшиеся предыдущими авторами, строгим методом в рамках теории случайных процессов, А. Я. Хинчин классифицировал схемы массового обслуживания. Следующим важным результатом явился метод вложенных цепей Маркова, созданный Д. Кендаллом и изложенный теперь почти во всех книгах по теории массового обслуживания (например, [14 и 19]). Данным методом решено большинство задач ТМО.

Для справок, кроме упомянутых выше работ, рекомендуются книги [6, 9, 12, 34, 36, 71].

В настоящее время в теории массового обслуживания рассматриваются следующие основные системы и направления: системы с приоритетным обслуживанием и с разделением времени; системы с ограничениями на длину очереди, на время пребывания требований в системе либо на время их ожидания; системы с групповым обслужива-

нием либо с групповым поступлением требований; системы с приборами ограниченной надежности; статистические методы восстановления характеристик систем по наблюдению различных случайных величин, связанных с процессом обслуживания.

Строгое математическое определение схем и задач теории массового обслуживания, которое удовлетворяло бы математиков и специалистов прикладных наук, в настоящее время вряд ли возможно. Это объясняется тем, что методы теории массового обслуживания тесно переплетаются с методами других разделов теории вероятностей и теории случайных процессов. Многие авторы предлагают обобщенные математические схемы для описания функционирования систем массового обслуживания. Повидному, процесс создания таких схем будет и впредь развиваться. Это направление представляется весьма перспективным, если учесть, что современные вычислительные машины дают возможность алгоритмизировать решение широких классов математических задач. Последнее относится не только к численным методам, но и к выводу формул.

Выбранный в книге математический аппарат в значительной мере предопределил и круг решаемых практических задач, связанных с созданием высоконадежных адаптивных систем управления летательными аппаратами, регламентацией их проверок, управлением посадкой групп самолетов, заправкой самолетов топливом в воздухе.

Гл. I (за исключением 1.1 и 1.3) написал И. Н. Коваленко; гл. II, III и 1.1 и 1.3 — Г. К. Москатов; гл. IV — Е. Ю. Барзилович; в 4.5 воспроизводится работа [6].

Авторы выражают глубокую признательность д-ру техн. наук, проф. Ю. И. Топчесеву за рекомендации по содержанию книги, а также благодарность чл.-корр. АН Литовской ССР, проф. Б. И. Григелионису за ценные замечания по содержанию первой главы книги.

Отзывы и замечания высылать по адресу: Москва, Б-78, 1-й Басманный пер., 3, издательство «Машиностроение».

Глава I

НЕКОТОРЫЕ ВОПРОСЫ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ. ОПТИМАЛЬНАЯ ОСТАНОВКА ПРОЦЕССА НАБЛЮДЕНИЙ

1.1. ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Рассмотрим основные понятия теории массового обслуживания в простейших предположениях, используемые в излагаемых ниже задачах проектирования систем управления ЛА.

Каналом обслуживания называется вся совокупность технических средств, а также людей, обеспечивающих обслуживание одной заявки. Функционирование каждого канала характеризуется временем, которое затрачивается на обслуживание одной заявки. Это время является случайным.

В простейших системах, описываемых марковскими процессами с конечным или счетным множеством состояний, рассмотрением которых мы ограничимся в данном параграфе, *время обслуживания* распределено по экспоненциальному закону с параметром μ . Это означает, что если в момент времени t_k происходит обслуживание заявки, но она еще не обслужена до конца, то закон распределения оставшегося времени обслуживания не зависит от того, сколько времени продолжалось обслуживание до момента t_k . Иногда говорят о «потоке обслуживаний», действующих на заявку, хотя каждая заявка может быть обслужена только один раз.

Будем считать, что:

- 1) все каналы системы имеют одинаковую интенсивность потока обслуживаний $\mu = (1/\bar{t}_{\text{обс}})$;
- 2) заявка может обслуживаться любым из n каналов, т. е. любой из n каналов доступен для заявки

Для правильного обслуживания необходимо знать то, что обслуживается, а также все черты чередования поступающих заявок. Поэтому изучение системы массового обслуживания (СМО) начинается с анализа структуры входящего потока заявок. При этом заявки отличаются друг от друга только моментами поступления на обслуживание.

Интенсивность стационарного пуассоновского потока заявок λ определяется средним интервалом времени \bar{t} между поступлением двух заявок:

$$\lambda = \frac{1}{\bar{t}} .$$

Эффективность (успешность) работы СМО зависит от параметров λ , μ , n и от дисциплины обслуживания, которая определяет порядок распределения заявок между свободными каналами: поведение заявок, попавших в

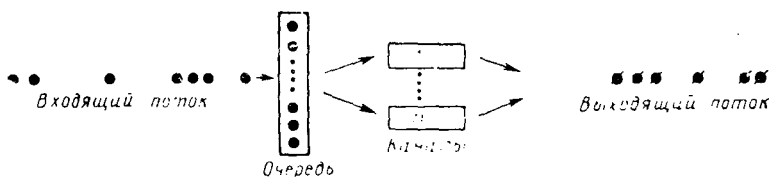


Рис. 1.1. Обобщенная структурная схема системы массового обслуживания с ожиданием

СМО на обслуживание, закон образования очереди, поведение заявки, попавшей в очередь, и т. п. [36].

Основная задача теории массового обслуживания — установление с необходимой точностью зависимости между характером потока заявок, числом каналов, их производительностью, дисциплиной обслуживания и эффективностью обслуживания.

Системы массового обслуживания можно разделить на *разомкнутые* и *замкнутые*. Отличительной чертой разомкнутой СМО является то, что источник потока заявок в систему не входит и его состояние не анализируется (рис. 1.1). Примером многоканальной разомкнутой системы является аэропорт с несколькими взлетно-посадочными полосами (ВПП) и входящим потоком взлетов и посадок самолетов.

В замкнутой СМО число источников заявок ограничено и интенсивность входящего потока зависит от состояния источников, обусловленного работой самой системы. Источник заявок является элементом СМО. Число клиентов такой системы ограничено числом каналов обслуживания. Пример — квазистационарная система автоматического регулирования. В период отработки рассогласования (обслуживания заявки) новые заявки не поступают.

В разомкнутой СМО каждое вновь поступившее требование, застав все каналы занятыми, становится на очередь. Если требование поступит в СМО, когда есть свободный канал, то оно сразу принимается на обслуживание.

Рассмотрим показатели эффективности СМО разомкнутого типа в установившемся режиме.

1. Приведенная плотность входящего потока:

$$\alpha = \frac{\lambda}{\mu} = \lambda \bar{t}_{\text{обс}}. \quad (1.1)$$

2. Вероятность того, что все обслуживающие каналы свободны:

$$P_0 = \left[\sum_{k=0}^{n-1} \frac{\alpha^k}{k!} + \frac{\alpha^n}{(n-1)!(n-\alpha)} \right]^{-1}, \quad (1.2)$$

где n — число каналов.

3. Вероятность того, что занято обслуживанием k каналов:

$$P_k = \frac{\alpha^k}{k!} P_0; \quad 1 \leq k < n. \quad (1.3)$$

4. Вероятность того, что все каналы заняты обслуживанием:

$$\pi = \frac{\alpha^n}{(n-1)!(n-\alpha)} P_0; \quad \frac{\alpha}{n} < 1. \quad (1.4)$$

5. Среднее время ожидания требованием начала обслуживания:

$$\bar{t}_{\text{ож}} = \frac{\pi \bar{t}_{\text{обс}}}{n-\alpha}; \quad \frac{\alpha}{n} < 1. \quad (1.5)$$

6. Средняя длина очереди:

$$\bar{v} = \lambda \bar{t}_{\text{ож}}. \quad (1.6)$$

7. Вероятность того, что все каналы заняты обслуживанием, и s требований находится в очереди:

$$P_{n+s} = \frac{\alpha^{n+s}}{n!n^s} P_0. \quad (1.7)$$

8. Вероятность того, что время пребывания требования в очереди больше некоторой величины t :

$$P(\tau > t) = \pi e^{-\nu(n-\alpha)t}. \quad (1.8)$$

9. Среднее число свободных каналов:

$$N_0 = \sum_{k=0}^{n-1} \frac{n-k}{k!} \alpha^k P_0. \quad (1.9)$$

10. Среднее число занятых каналов:

$$N_3 = n - N_0. \quad (1.10)$$

11. Коэффициент занятости:

$$K_3 = \frac{N_3}{n}. \quad (1.11)$$

12. Коэффициент простоя:

$$K_n = \frac{N_0}{n}. \quad (1.12)$$

В авиационной практике имеют место случаи, когда на время пребывания заявки в СМО или на среднюю длину очереди наложены ограничения («нетерпеливая» заявка). В таких СМО, называемых системами *смешанного типа*, «нетерпеливая» заявка, не дождавшись обслуживания, может покинуть систему.

Приведем формулы для расчета основных характеристик многоканальной СМО смешанного типа в стационарном режиме.

1. Вероятность того, что все каналы системы свободны от обслуживания:

$$P_0 = \left[\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{m=1}^s (n+m\beta)} \right]^{-1}, \quad (1.13)$$

где $\alpha = \lambda \bar{t}_{обс}$; $\beta = \frac{\bar{t}_{обс}}{\bar{t}_{ож}}$.

2. Вероятность того, что заняты обслуживанием k каналов системы:

$$P_k = \frac{\frac{\alpha^k}{k!}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{m=1}^s (n + \beta m)}}; \quad 0 \leq k \leq n. \quad (1.14)$$

3. Вероятность того, что обслуживанием требований заняты все n каналов системы и s требований ожидают обслуживания:

$$P_{n+s} = \frac{\frac{\alpha^{n+s}}{n! \prod_{m=1}^s (n + \beta m)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{m=1}^s (n + \beta m)}}; \quad s \geq 1. \quad (1.15)$$

4. Среднее число требований, ожидающих обслуживания:

$$\bar{v} = \sum_{s=1}^{\infty} s P_{n+s} = \frac{\frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{s \alpha^s}{\prod_{m=1}^s (n + \beta m)}}{\sum_{k=0}^n \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{s=1}^{\infty} \frac{\alpha^s}{\prod_{m=1}^s (n + \beta m)}}. \quad (1.16)$$

5. Вероятность отказа в обслуживании требования:

$$P_{\text{отк}} = \frac{\beta}{\alpha} \bar{v}. \quad (1.17)$$

6. Среднее число каналов, занятых обслуживанием:

$$N_3 = \sum_{k=1}^n k P_k + n \sum_{s=1}^{\infty} P_{n+s}. \quad (1.18)$$

7. Коэффициент загрузки каналов обслуживания:

$$K_3 = \frac{N_3}{n}. \quad (1.19)$$

8. Среднее число свободных каналов обслуживания:

$$N_0 = n - N_3 = \sum_{k=0}^n (n-k) P_k. \quad (1.20)$$

9. Коэффициент простоя каналов обслуживания:

$$K_n = \frac{N_0}{n}. \quad (1.21)$$

10. Вероятность того, что требование, прибывшее в систему, будет обслужено:

$$P_{обс} = 1 - P_{отк}. \quad (1.22)$$

В целях облегчения вычислений составлены таблицы для определения P_k и $P_{отк}$ при $k \leq n$ с тремя входами α , β , n (см. [35]).

Для приближенных расчетов можно заменить бесконечные суммы их приближенными значениями:

$$\sum_{s=r}^{\infty} \frac{\alpha^s}{\prod_{m=1}^{\infty} (n + m\beta)} < \frac{\left(\frac{\alpha}{\beta}\right)^r}{r!} e^{\frac{\alpha}{\beta}}; \quad (1.23)$$

$$\sum_{s=r}^{\infty} \frac{s^{\alpha s}}{\prod_{m=1}^{\infty} (n + m\beta)} < \frac{\left(\frac{\alpha}{\beta}\right)^r}{(r-1)!} e^{\frac{\alpha}{\beta}}. \quad (1.24)$$

Формулы (1.13) — (1.24) получены в предположении, что время обслуживания требований и их ожидания в очереди распределены по показательному закону. Статистическое моделирование СМО смешанного типа в стационарном режиме при различных законах распределения случайного времени обслуживания и ожидания [35] свидетельствует о том, что эти зависимости приближенно выполняются и для других законов. Для иллюстрации приведем таблицу расчетов вероятности отказа в обслуживании $P_{отк}$ методом Монте-Карло (табл. 1.1).

Таблица 1.1

Закон распределения $\bar{t}_{\text{обс}}$ и $\bar{t}_{\text{ож}}$	Варианты при $n = 1$				Варианты при $n = 3$			
	$\alpha=7$ $\beta=1.5$	$\alpha=1,5$ $\beta=0,33$	$\alpha=6$ $\beta=0,5$	$\alpha=20$ $\beta=2$	$\alpha=3$ $\beta=2$	$\alpha=20$ $\beta=0.5$	$\alpha=10$ $\beta=1$	$\alpha=15$ $\beta=1$
Показательный . . .	0,861	0,386	0,788	0,949	0,226	0,783	0,675	0,784
Релея . . .	0,863	0,392	0,791	0,948	0,228	0,786	0,672	0,781
Нормальный усеченный	0,859	0,391	0,787	0,953	0,224	0,785	0,674	0,780
Равномерный	0,857	0,393	0,793	0,951	0,226	0,787	0,673	0,783
По формуле (1.17) . .	0,860	0,390	0,790	0,950	0,225	0,785	0,673	0,785

Из таблицы видно, сколь малыми оказываются погрешности от замены экспоненциальным законом законов распределения Релея, усеченного нормального и равномерного [35].

Подобное обстоятельство имеет место далеко не для всех СМО и функционалов, характеризующих их эффективность.

В качестве примера приведем следующий результат. Пусть имеется резервированная система, т. е. такая, для которой выход из строя одного элемента не приводит к отказу системы. Задастся множество отказовых состояний системы R . Допустим, что попадание в множество R возможно лишь при условии, что имеются неисправные элементы, ожидающие восстановления (из-за занятости восстанавливающих каналов). Пусть λ_{ij} — интенсивность отказов элементов в различных состояниях, $H_{ij}(x)$ — функции распределения времени восстановления элементов в различных состояниях, $F(x)$ — функция распределения времени вхождения системы в множество состояний R , $G(x)$ — аналогичная функция распределения для случая, когда $H_{ij}(x)$ заменены экспонентами с теми же математическими ожиданиями. Тогда параметры можно подобрать таким образом, что G будет сколь угодно близко к $1 - e^{-x}$. В то же время $F(x)$ будет сколь угодно близко к единице или к нулю по нашему желанию.

Третий класс систем, который будет рассматриваться в дальнейшем, — это СМО с *приоритетным обслуживанием*. В таких системах могут быть различные варианты дисциплины обслуживания. Системами массового обслуж-

живания с *абсолютным* приоритетом называются такие системы, в которых заявка, обладающая приоритетом, немедленно принимается к обслуживанию каналом, занятым обслуживанием заявки без приоритета в обслуживании. После того, как требование, обладающее приоритетом, будет обслужено и других требований, обладающих приоритетом, нет, возобновляется прерванное обслуживание требования, не обладающего приоритетом. При этом возможны различные варианты:

- требование, обслуживание которого прервано, начинает обслуживаться заново;
- прерванное обслуживание требования начинается с того места, где оно было прервано;
- требование, обслуживание которого было прервано, вообще теряется.

Системами массового обслуживания с *относительным* приоритетом называются такие системы, в которых требование, не обладающее приоритетом, обслуживается до конца, после чего принимаются к обслуживанию требования, обладающие приоритетом [36].

Мы рассмотрим наиболее простую систему из этого класса, а именно, одноканальную СМО с абсолютным приоритетом. На вход системы поступают два независимых простейших потока требований с интенсивностями λ_1 и λ_2 . Требования первого потока с интенсивностью λ_1 обладают приоритетом в обслуживании. Длина очереди для требований обоих видов не ограничена. Если канал обслуживает заявку первого потока, то интенсивность простейшего потока обслуживаний равна μ_1 . Требования второго потока обслуживаются с интенсивностью μ_2 .

При сделанных допущениях нет различия между следующими двумя дисциплинами обслуживания:

- прерванное обслуживание требования начинается с того момента, где оно было прервано;
- требование, обслуживание которого было прервано, начинает обслуживаться заново.

Показатели эффективности функционирования такой СМО в стационарном режиме определяются по приводимым ниже формулам.

1. Среднее число приоритетных требований, находящихся в очереди:

$$\bar{v}_1 = \frac{(\alpha_1)^2}{1 - \alpha_1}, \quad (1.25)$$

где $\alpha_1 = (\lambda_1/\mu_1)$.

2. Среднее время пребывания в очереди требования, обладающего приоритетом:

$$\bar{t}_{ож_1} = \frac{\bar{v}_1}{\lambda_1} = \frac{1}{\mu_1} \frac{\alpha_1}{1 - \alpha_1}. \quad (1.26)$$

3. Среднее время пребывания требования, обладающего приоритетом, в системе (в очереди и на обслуживании):

$$\bar{t}_{пр} = \bar{t}_{ож_1} + \bar{t}_{обс_1} = \frac{1}{\mu_1} \frac{1}{1 - \alpha_1}. \quad (1.27)$$

4. Среднее число неприоритетных требований, находящихся в системе:

$$\bar{v}_2 = \frac{\alpha_2}{1 - \alpha} \left[1 + \frac{\mu_2}{\mu_1} \frac{\alpha_1}{1 - \alpha_1} \right], \quad (1.28)$$

где $\alpha = \alpha_1 + \alpha_2$.

При $\mu_1 = \mu_2$

$$\bar{v}_2 = \frac{\alpha_1}{(1 - 2\alpha_1)(1 - \alpha_1)}. \quad (1.29)$$

5. Среднее время нахождения в системе требования, не обладающего приоритетом:

$$\bar{t}_{пр_2} = \bar{t}_{ож_2} + \bar{t}_{обс_2} = \bar{t}_{ож_2} + \frac{1}{\mu_2}. \quad (1.30)$$

6. Среднее время ожидания требования, не обладающего приоритетом:

$$\bar{t}_{ож_2} = \frac{1}{\mu_2} \frac{\frac{\mu_2}{\mu_1} \frac{\alpha_1}{1 - \alpha} + \alpha}{1 - \alpha}. \quad (1.31)$$

В заключение укажем основные причины, в силу которых имеет практический интерес теория массового обслуживания в простейших предпосылках.

1. Замена пуассоновских потоков событий пуассоновскими с теми же интенсивностями приводит в большинстве практических задач к решениям, столь мало отличающимся от истинных, что этой погрешностью (3—5%) можно пренебречь. Лишь в редких случаях погрешность доходит до 10—12%, что приемлемо при решении ряда задач проектирования систем¹. Это объясняется

¹ При решении сложных задач, когда нет уверенности в том, что замена реальных потоков пуассоновскими потоками приведет к малым ошибкам, рекомендуется проверять аналитическое решение методом Монте-Карло [36].

тем, что потоки событий, протекающие в реальных СМО, в силу предельных теорем теории потоков весьма близки к пуассоновским.

2. Природа процессов во многих СМО такова, что к простейшему входящему потоку при одинаковых плотностях им приспособиться труднее. В ряде случаев можно показать [14], что если СМО рассчитывается на этот случай, то обслуживание других случайных потоков с одинаковой плотностью поступления заявок будет «надежнее».

3. Если не делать предположения о том, что процесс, протекающий в СМО, является марковским, то для его аналитического исследования потребуется значительно более сложный математический аппарат (например, система интегро-дифференциальных уравнений с переменными коэффициентами).

1.2. МАРКОВСКИЕ ПРОЦЕССЫ С КОНЕЧНЫМ ИЛИ СЧЕТНЫМ МНОЖЕСТВОМ СОСТОЯНИЙ

К структуре марковских процессов с конечным или счетным¹ множеством состояний мы подойдем, рассмотрев вначале следующий простой пример.

Пусть η положительная случайная величина с функцией распределения $F(x)$ и плотностью $p(x)$. Определим случайный процесс $v(t)$ ($t \geq 0$) такой формулой:

$$v(t) = \begin{cases} 0 & \text{при } t \leq \eta; \\ 1 & \text{при } t > \eta. \end{cases} \quad (1.32)$$

Траектории этого процесса будут ступенчатыми функциями, каждая из которых в некоторой точке имеет скачок, равный единице. Обратим внимание на следующие свойства случайного процесса $v(t)$.

1. Если задано значение $v(t_0) = i$, то поведение процесса $v(t)$ после момента t_0 не зависит от значений $v(t)$ при $t < t_0$.

2. Если задано значение процесса $v(t) = 0$, то за малое время $h > 0$ процесс может перейти в состояние «1» с вероятностью $\lambda(t)h + o(h)$, где $\lambda(t)$ — некоторая неотрицательная функция; если $\lambda(t) = 1$, то $v(t+x) = 1$ при любых $x > 0$.

¹ Счетным множеством называется бесконечное множество, элементы которого можно пронумеровать числами натурального ряда: 1, 2, ..., n , Так, например, множество пар (n_1, n_2) , троек (n_1, n_2, n_3) , и т. д., где n_i — целые числа, является счетным.

3. Никакие изменения состояния $v(t)$, кроме перехода из состояния «0» в состояние «1» невозможны.

Проверим первое свойство. Если $v(t_0) = 0$, то, очевидно, $v(t) = 0$ при всех $t < t_0$. Это означает, что при данном условии величины $v(t)$ ($t < t_0$) являются детерминированными. Отсюда следует, что поведение процесса $v(t)$ при $t > t_0$ от них не зависит. (Известно, что любая случайная величина или случайный процесс не зависят от любой детерминированной величины). Пусть теперь $v(t_0) = 1$. Тогда $v(t) = 1$ при всех $t > t_0$, независимо от того, при каких $t < t_0$ процесс равнялся нулю, а при каких — единице.

Сформулированное свойство называется *марковским свойством случайного процесса*.

Проверим второе свойство. Предположим, что $v(t) = 0$ и возьмем некоторое $h > 0$. Событие $\{v(t+h) = 0\}$ равносильно событию $\{\xi > t+h\}$, а событие $\{v(t) = 0\}$ — событию $\{\xi > t\}$. Событие же $\{v(t+h) = 1\}$ равносильно событию $\{\xi \leq t+h\}$. Тогда

$$\begin{aligned} P(v(t+h) = 1 | v(t) = 0) &= P(\xi \leq t+h | \xi > t) = \\ &= P(t < \xi \leq t+h | \xi > t) = \frac{P(t < \xi \leq t+h)}{P(\xi > t)} = \\ &= \frac{F(t+h) - F(t)}{1 - F(t)} = \frac{p(t)}{1 - F(t)} h + o(h) = \lambda(t) h + o(h), \end{aligned} \quad (1.33)$$

если обозначить

$$\lambda(t) = \frac{p(t)}{1 - F(t)}. \quad (1.34)$$

Итак, оба свойства, сформулированные выше, выполняются.

Заметим еще, что $\lambda(t)$ неотрицательна, непрерывна и

$$\int_0^{\infty} \lambda(t) dt = \infty. \quad \text{Действительно,}$$

$$\lambda(t) = \frac{F'(t)}{1 - F(t)} = - \frac{[1 - F(t)]'}{1 - F(t)} = - [\ln [1 - F(t)]]'.$$

Отсюда

$$\int_0^A \lambda(t) dt = \ln [1 - F(0)] - \ln [1 - F(A)]. \quad (1.35)$$

Так как при $A \rightarrow \infty$ $F(A) \rightarrow 1$ и $\ln [1 - F(A)] \rightarrow -\infty$, то

$$\int_0^{\infty} \lambda(t) dt = \infty. \quad (1.36)$$

Поставим теперь вопрос: однозначно ли первое и второе свойства определяют случайный процесс (в статистическом смысле)? Не будем учитывать существование случайной величины ξ , определяющей процесс $v(t)$, и предположим, что существует случайный процесс $v(t)$, удовлетворяющий первому и второму свойствам, а также условию $v(0) = 0$, причем $\lambda(t)$ — неотрицательная непрерывная функция с бесконечным интегралом. Обозначим через $\Phi(t)$ вероятность события $\{v(t) = 0\}$. Рассмотрим два момента времени: t и $t+h$, где $h > 0$ — малое число. Ясно, что если $v(t+h) = 0$, то и $v(t) = 0$. Поэтому

$$P(v(t+h) = 0) = P(v(t) = 0) P(v(t+h) = 0 | v(t) = 0) \quad (1.37)$$

или, вследствие второго условия

$$P(v(t+h) = 0) = P(v(t) = 0) [1 - \lambda(t)h + o(h)], \quad (1.38)$$

т. е.

$$\Phi(t+h) = \Phi(t) [1 - \lambda(t)h + o(h)]. \quad (1.39)$$

Последнее равенство можно переписать так:

$$\frac{\Phi(t+h) - \Phi(t)}{h} = -\lambda(t)\Phi(t) + o(1). \quad (1.40)$$

Приняв $h \rightarrow 0$, найдем, что функция $\Phi(t)$ обладает правосторонней производной $\Phi'(t)$, удовлетворяющей уравнению

$$\Phi'(t) = -\lambda(t)\Phi(t). \quad (1.41)$$

Можно показать, что это уравнение сохраняет смысл и для двусторонней производной функции $\Phi(t)$, но тогда

$$\frac{d \ln \Phi(t)}{dt} = -\lambda(t),$$

откуда

$$\ln \Phi(t) - \ln \Phi(0) = -\int_0^t \lambda(x) dx. \quad (1.42)$$

Вспомним, что $\Phi(t) = P(v(t) = 0)$; отсюда $\Phi(0) = 1$ и

$$\Phi(t) = \exp \left\{ - \int_0^t \lambda(x) dx \right\}. \quad (1.43)$$

Теперь обозначим через ξ момент времени, когда процесс $v(t)$ испытывает скачок, равный единице. Если $F(t)$ — функция распределения случайной величины ξ , то

$$\begin{aligned} F(t) &= P(\xi < t) = P(v(t) = 1) = 1 - P(v(t) = 0) = \\ &= 1 - \Phi(t) = 1 - \exp \left\{ - \int_0^t \lambda(x) dx \right\}. \end{aligned} \quad (1.44)$$

Продифференцировав это выражение по t , найдем, что случайная величина ξ обладает плотностью

$$p(t) = [1 - F(t)] \lambda(t). \quad (1.45)$$

Итак, исходя из рассмотренных выше свойств случайного процесса, мы выразили его через некоторую случайную величину ξ так, что $v(t) = 0$ при $t \leq \xi$ и $v(t) = 1$ при $t > \xi$ ¹.

Если $\int_0^{\infty} \lambda(t) dt < \infty$, тогда величину ξ приходится интерпретировать как величину, зависящую от случая и способную принимать значение ∞ с положительной вероятностью

$$P(\xi = \infty) = \exp \left\{ - \int_0^{\infty} \lambda(t) dt \right\}. \quad (1.46)$$

При конечных t , как и прежде

$$P(\xi > t) = \exp \left\{ - \int_0^t \lambda(x) dx \right\}. \quad (1.47)$$

Определим теперь марковский процесс с конечным или счетным множеством состояний. Это будет некоторое построение, в котором процессы типа, рассмотренного выше, будут играть роль составных элементов.

Пусть Ω — множество состояний марковского процесса. Определим две системы функций: $\lambda_i(t), i \in \Omega$ и $\pi_{ij}(t)$,

¹ Строго говоря, непрерывность $v(t)$ слева не следует из рассмотренных выше условий и свойств $\lambda(t)$ и вводится как дополнительное условие. При применении это не может вызвать неоднозначного толкования.

$i \in \Omega, j \in \Omega, i \neq j$. Потребуем, чтобы все эти функции были неотрицательными и, кроме того,

$$\sum_{j \neq i} \pi_{ij}(t) = 1.$$

Таким образом, при фиксированных t и t числа $\pi_{ij}(t)$ можно рассматривать как задающие распределение вероятностей на множестве Ω .

Марковский процесс $v(t)$ с множеством состояний Ω определяется следующими свойствами.

1. Если задано значение $v(t_0) = i$, то поведение процесса $v(t)$ после момента t_0 не зависит от значений $v(t)$ при $t < t_0$.

2. При малых $h > 0$

$$P(v(t+h) = j | v(t) = i) = \lambda_i(t) \pi_{ij}(t) h + o(h); \quad j \neq i.$$

3. При малых $h > 0$

$$P(v(t+\tau) = i; 0 \leq \tau \leq h | v(t) = i) = 1 - \lambda_i(t) h + o(h).$$

4. Значение процесса $v(t)$ в момент $t=0$ задается распределением вероятностей

$$P(v(0) = i) = p_i^{(0)}; \quad i \in \Omega,$$

где $\sum_{i \in \Omega} p_i^{(0)} = 1$.

Функция $\lambda_i(t)$ называется мгновенной интенсивностью выхода процесса $v(t)$ из i -го состояния в момент t ; функция $\lambda_{ij}(t) = \lambda_i(t) \pi_{ij}(t)$ называется мгновенной интенсивностью перехода процесса $v(t)$ из состояния i в состояние j , при этом $\pi_{ij}(t)$ есть условная вероятность попадания процесса в состояние j при условии, что в момент t произошел выход из состояния i .

Вероятностную структуру марковского процесса $v(t)$ можно разъяснить следующим образом.

Состояние процесса при $t=0$ определяется вероятностями $p_i^{(0)}$, а именно $P(v(0) = i) = p_i^{(0)}$, $i \in \Omega$. Траектория процесса $v(t)$ при $t \geq 0$ есть ступенчатая функция, скачки которой происходят в моменты t_1, t_2, \dots ($0 \leq t_1 < t_2 < \dots$). Если $v(0) = i$, то t_1 есть случайная величина с функцией распределения $P(t_1 < t) = 1 - \exp \left\{ - \int_0^t \lambda_i(x) dx \right\}$.

В момент t_1 процесс переходит в состояние $j \neq i$ с вероятностью $\pi_{ij}(t_1)$.

Допустим, что в момент $t_n (n \geq 1)$ процесс $v(t)$ перешел в состояние i . Тогда t_{n+1} определялся как случайная величина с функцией распределения

$$P(t_{n+1} < t) = 1 - \exp \left\{ - \int_{t_n}^t \lambda_i(x) dx \right\} \quad (t \geq t_n). \quad (1.48)$$

В момент t_{n+1} процесс переходит в состояние $j \in \Omega$ с вероятностью $\pi_{ij}(t_{n+1})$. При этом, вообще говоря, может случиться, что за конечный интервал времени накопится бесконечное множество скачков, так что описанная конструкция определит процесс лишь до первого момента накопления скачков. Однако с прикладной точки зрения такая ситуация является весьма исключительной. Достаточным условием того, что $t_n \rightarrow \infty$ при $n \rightarrow \infty$, т. е. условием, при котором отрезки $[t_n, t_{n+1}]$ покрывают всю полу-прямую $[0, \infty]$, является неравенство

$$\lambda_i(t) \leq C; \quad i \in \Omega.$$

Основные характеристики марковского процесса с конечным или счетным множеством состояний

Пусть $p_j(t)$ — вероятность того, что $v(t) = j$ ($j \in \Omega$). Можно записать

$$p_j(t) = \sum_{i \in \Omega} p_i^{(0)} p_{ij}(t), \quad (1.49)$$

где $p_{ij}(t) = P(v(t) = j | v(0) = i)$.

Таким образом, $p_j(t)$ можно выразить через функции $p_{ij}(t)$, называемые переходными вероятностями марковского процесса.

Рассмотрев близкие моменты времени t и $t+h$ и используя определяющие свойства марковского процесса, найдем

$$p_{ij}(t+h) = p_{ij}(t) [1 - \lambda_j(t)h + o(h)] + \sum_{k \neq j} p_{ik}(t) \lambda_{kj}(t) h + o(h).$$

Отсюда

$$\begin{aligned} \frac{1}{h} [p_{ij}(t+h) - p_{ij}(t)] &= -\lambda_j(t) p_{ij}(t) + \\ &+ \sum_{k \neq j} p_{ik}(t) \lambda_{kj}(t) + o(1). \end{aligned}$$

При $h \rightarrow 0$

$$p'_{ij}(t) = -\lambda_j(t) p_{ij}(t) + \sum_{k \neq j} p_{ik}(t) \lambda_{kj}(t); \quad i \in \Omega; \quad j \in \Omega. \quad (1.50)$$

В нашем рассуждении было существенно, что $h > 0$, однако, можно показать, что $p_{ij}(t)$ обладает также двусторонней производной.

Уравнения (1.50) принято называть уравнениями Колмогорова. Система этих уравнений при фиксированном i имеет единственное решение при начальных условиях:

$$p_{ii}(0) = 1; \quad p_{ij}(0) = 0; \quad j \neq i,$$

по крайней мере в случае, когда $\lambda_j(t) \leq C, j \in \Omega$.

Часто интересно иметь явный вид вероятностей $p_{ij}(t)$. Это можно вывести с помощью следующего рассуждения. По формуле полной вероятности

$$p_{ij}(t) = \sum_{m=0}^{\infty} p_{ij}^{(m)}(t), \quad (1.51)$$

где $p_{ij}^{(m)}(t)$ — вероятность события, состоящего в том, что $v(t) = j$ и в интервале $(0; t)$ произошло m скачков процесса $v(t)$ при условии, что $v(0) = i$.

Функции, служащие членами выписанного ряда, можно рассчитывать по простой рекуррентной формуле. Однако, чтобы сделать это, необходимо ввести несколько более общие функции. Именно, обозначим через $p_{ij}^{(m)}(t, \tau)$ ($t \leq \tau$) условную вероятность события, состоящего в том, что $v(\tau) = j$ и в интервале (t, τ) произошло m скачков процесса $v(t)$ при условии $v(t) = i$. Очевидно, $p_{ij}^{(m)}(t) = p_{ij}^{(m)}(0, t)$.

Введенные функции удовлетворяют следующим соотношениям:

при $m \geq 0$

$$p_{ij}^{(0)}(t, \tau) = p_{ij}(0) \exp \left\{ - \int_t^{\tau} \lambda_i(x) dx \right\}; \quad (1.52)$$

при $m \geq 1$

$$p_{ij}^{(m)}(t, \tau) = \sum_{k \neq j} \int_t^{\tau} p_{ik}^{(m-1)}(u) \lambda_{kj}(u) \exp \left\{ - \int_u^{\tau} \lambda_j(x) dx \right\} du. \quad (1.53)$$

¹ Здесь используется условие, что $t_n \rightarrow \infty$ при $n \rightarrow \infty$.

Данные формулы особенно удобны для расчетов в случае, если интенсивности переходов можно рассматривать как малые величины. Приведем одну простую оценку. Предположим, что для всех $i \in \Omega$ выполняется оценка

$$\int_0^t \lambda_i(x) dx \leq a$$

для некоторого фиксированного $t > 0$. Тогда справедливо неравенство

$$\sum_{j \in \Omega} \left[p_{ij}(t) - \sum_{m=0}^r p_{ij}^{(m)}(t) \right] \leq e^{-a} \sum_{m=r+1}^{\infty} \frac{a^m}{m!}; \quad r = 0, 1, 2, \dots \quad (1.54)$$

В некоторых случаях полезно представить $\lambda_i(t)$ в виде функции некоторого малого параметра ε и найти разложение вероятностей $p_{ij}(t, \tau)$ по степеням этого параметра. Это делается следующим образом. Будем писать $\varepsilon \lambda_i(t)$ вместо $\lambda_i(t)$, чтобы не вводить новых обозначений. (Иначе говоря, мы считаем, что $\lambda_i(t) = \varepsilon \lambda_i^{(0)}(t)$, где $\varepsilon > 0$ — малый параметр, $\lambda_i^{(0)}(t)$ — фиксированные функции; затем, для краткости, $\lambda_i^{(0)}(t)$ обозначаем через $\lambda_i(t)$. Подобным же образом будем писать $\varepsilon \lambda_{ij}(t)$ вместо $\lambda_{ij}(t)$; $\pi_{ij}(t)$ считаем не зависящим от ε .

С учетом сказанного найденные выше уравнения перепишутся следующим образом:

при $m = 0$

$$p_{ij}^{(0)}(t, \tau) = p_{ij}(0) \exp \left\{ -\varepsilon \int_t^{\tau} \lambda_i(x) dx \right\}; \quad (1.55)$$

при $m \geq 1$

$$p_{ij}^{(m)}(t, \tau) = \sum_{k \in \Omega} \varepsilon \int_t^{\tau} p_{ik}^{(m-1)}(u) \lambda_{kj}(u) \exp \left\{ -\varepsilon \int_u^{\tau} \lambda_j(x) dx \right\} du. \quad (1.56)$$

Обозначим

$$\Lambda_j(t, \tau) = \int_t^{\tau} \lambda_j(x) dx; \quad t \leq \tau, \quad j \in \Omega. \quad (1.57)$$

Тогда

$$\exp \left\{ -\varepsilon \int_0^{\tau} \lambda_j(x) dx \right\} = \sum_{s=0}^{\infty} \frac{(-1)^s \varepsilon^s}{s!} \Lambda_j^s(t, \tau). \quad (1.58)$$

Рассмотрим ряды вида

$$p_{ij}^{(m)}(t, \tau) = \sum_{s=m}^{\infty} \varepsilon^s p_{ij}^{(m,s)}(t, \tau). \quad (1.59)$$

Подставив это неравенство в рекуррентные формулы, найдем: при $m=0$

$$\sum_{s=0}^{\infty} \varepsilon^s p_{ij}^{(0,s)}(t, \tau) = p_{ij}(0) \sum_{s=0}^{\infty} \frac{(-1)^s \varepsilon^s}{s!} \Lambda_j^s(t, \tau); \quad (1.60)$$

при $m \geq 1$

$$\begin{aligned} & \sum_{s=m}^{\infty} \varepsilon^s p_{ij}^{(m,s)}(t, \tau) = \\ & = \sum_{k \neq j} \varepsilon \int_0^{\tau} \sum_{s=m-1}^{\infty} \varepsilon^s p_{ik}^{(m-1,s)}(u) \lambda_{kj}(u) \sum_{r=0}^{\infty} \frac{(-1)^r \varepsilon^r}{r!} \Lambda_j^r(u, \tau) du. \end{aligned} \quad (1.61)$$

Коэффициенты разложения найдем, приравнявая коэффициенты при одинаковых степенях ε в левой и правой частях:

при $m=0$

$$p_{ij}^{(0,s)}(t, \tau) = p_{ij}(0) \frac{(-1)^s}{s!} \Lambda_j^s(t, \tau); \quad (1.62)$$

при $m \geq 1$

$$\begin{aligned} & p_{ij}^{(m,s)}(t, \tau) = \\ & = \sum_{k \neq j} \sum_{r=0}^{s-m} \frac{(-1)^r}{r!} \int_0^{\tau} p_{ik}^{(m-1, s-r-1)}(u) \lambda_{kj}(u) \Lambda_j^r(u, \tau) du; \quad s \geq m. \end{aligned} \quad (1.63)$$

Рассмотрим теперь ряд ¹

$$p_{ij}(t, \tau) = \sum_{s=0}^{\infty} \varepsilon^s p_{ij}^{(s)}(t, \tau). \quad (1.64)$$

¹ Обозначение «(·^s)» показывает, что по m произведено суммирование.

Тогда

$$p_{ij}^{(s)} = \sum_{m=0}^s p_{ij}^{(m,s)}; \quad s=0,1,2,\dots, \quad (1.65)$$

а формулы для слагаемых последней суммы приведены выше.

Пусть A — некоторое множество состояний процесса $v(t): A \in \Omega$. Обозначим через $P_i^A(t)$ вероятность того, что $v(t) \in A$ для всех τ ($0 \leq \tau \leq t$) при условии, что $v(0) = i$.

Вычисление вероятностей $P_i^A(t)$ легко сводится к рассмотренной выше задаче. Введем вспомогательный марковский процесс $\bar{v}(t)$, характеристики которого задаются формулой

$$\bar{\lambda}_{ij}(t) = \begin{cases} \lambda_{ij}(t), & \text{если } i \in A, j \in A; \\ 0 & \text{в противном случае.} \end{cases} \quad (1.66)$$

Если для такого процесса найти вероятности $\bar{p}_{ij}(t)$ перехода из состояния i в момент $t=0$ и в состояние j в момент t , то

$$P_i^A(t) = \sum_{j \in A} \bar{p}_{ij}(t). \quad (1.67)$$

Из этой формулы следует, что

$$P_i^A(t) = 0; \quad i \notin A. \quad (1.68)$$

Пусть $v(t)$ — марковский процесс с интенсивностями перехода $\lambda_{ij}(t)$ и распределением вероятностей $p_i^{(0)}$, $i \in \Omega$ состояний в момент $t=0$. Предположим далее, что $f(i)$, $i \in \Omega$ — некоторая числовая функция, действительная или комплексная. Представляет интерес рассмотреть функционалы вида

$$\eta = \eta_T = \int_0^T f(v(t)) dt.$$

Так, положив $f(i) = 1$ при $i \in A$, $f(i) = 0$ при $i \notin A$, получим, что η_T есть суммарное время пребывания процесса $v(t)$ в множестве состояний A в течение интервала $(0; T)$.

Покажем, как найти $M[\eta]$ и $M[\eta^2]$, если они существуют¹.

¹ Здесь и далее символом $M[\dots]$ обозначено математическое ожидание случайной величины или функции.

Имеем:

$$\left. \begin{aligned}
 M[\eta] &= \int_0^T M[f(v(t))] dt = \sum_{j \in \Omega} f(j) \int_0^T p_j(t) dt; \\
 M[\eta^2] &= \int_0^T \int_0^T M[f(v(t))f(v(\tau))] dt d\tau = \\
 &= 2 \iint_{0 < t < \tau < T} M[f(v(t))f(v(\tau))] dt d\tau = \\
 &= 2 \sum_{i, j \in \Omega} f(i)f(j) \int_0^T \int_0^T P(v(t)=i, v(\tau)=j) dt d\tau = \\
 &= 2 \sum_{i, j \in \Omega} f(i)f(j) \int_0^T \int_0^T p_i(t) p_{ij}(t, \tau) dt d\tau.
 \end{aligned} \right\} (1.69)$$

Однородные марковские процессы с конечным или счетным множеством состояний

Указанным термином определяют марковские процессы с конечным или счетным множеством состояний, для которых интенсивности перехода $\lambda_{ij}(t)$ не зависят от t :

$$\lambda_{ij}(t) = \lambda_{ij}; \quad i \in \Omega; \quad j \in \Omega; \quad i \neq j; \quad t > 0. \quad (1.70)$$

Выпишем уравнения для переходных вероятностей однородного марковского процесса:

$$p'_{ij}(t) = -\lambda_j p_{ij}(t) + \sum_{k \neq j} \lambda_{kj} p_{ik}(t); \quad i \in \Omega; \quad j \in \Omega, \quad (1.71)$$

где $\lambda_j = \sum_{k \neq j} \lambda_{jk}$.

В данном случае это система уравнений с постоянными коэффициентами и ее можно решить методом преобразований Лапласа.

Если при всех $i \in \Omega$ вероятность $p_{ij}(t)_{t \rightarrow \infty} \rightarrow p_j$, где p_j — распределение вероятностей на Ω , то говорят, что марковский процесс $v(t)$ обладает эргодическим распределением p_j , а сам процесс называют эргодическим.

Свойство эргодичности можно объяснить так. Вероятность застать процесс $v(t)$ в некотором состоянии в момент t стремится при $t \rightarrow \infty$ к некоторой вероятности, не зависящей от состояния процесса в момент $t=0$. Если

процесс обладает эргодическим распределением, то последнее определяется системой уравнений:

$$\lambda_j p_j = \sum_{k \neq j} \lambda_{kj} p_k; \quad j \in \Omega \quad (1.72)$$

с нормирующим условием $\sum_{j \in \Omega} p_j = 1$.

Рассмотрим для простоты однородный марковский процесс $v(t)$ с конечным множеством состояний. Скажем, что состояние j достижимо из состояния i , если $\lambda_{ij} > 0$, либо существует цепочка состояний i_1, i_2, \dots, i_m такая, что $\lambda_{i i_1} \lambda_{i_1 i_2} \lambda_{i_2 i_3} \dots \lambda_{i_m j} > 0$.

Если j достижимо из i и i достижимо из j , скажем, что i и j — сообщающиеся состояния. Класс состояний C назовем замкнутым, если из любого состояния внутри C не достижимы состояния вне C . Признаком замкнутости класса состояний C служит равенство $\lambda_{ij} = 0$ для всех $i \in C, j \notin C$.

Во множестве Ω состояний процесса $v(t)$ выделится конечное число замкнутых классов C_1, \dots, C_r , сообщающихся состояний, называемых эргодическими классами; состояния, не попавшие ни в один из классов C_1, \dots, C_r , называются несущественными состояниями. Если i некоторое состояние, принадлежащее классу $C_s (1 \leq s \leq r)$, то это состояние сообщается с каждым из состояний, принадлежащих тому же классу C_s , и не сообщается ни с одним состоянием, не принадлежащим C_s . Из состояния i не достижимо ни одно состояние, не принадлежащее C_s .

Если процесс в некоторый момент времени попал в один из эргодических классов состояний, он навсегда в нем и останется. С другой стороны, в каком бы состоянии процесс ни находился при $t=0$, он с вероятностью, равной единице, в некоторый момент попадает в какой-нибудь из эргодических классов состояний.

Рассмотрим систему уравнений:

$$\left. \begin{aligned} \lambda_j p_j &= \sum_{k \neq j} \lambda_{kj} p_k; \quad j \in C_s; \\ \sum_{j \in C_s} p_j &= 1 \end{aligned} \right\} \quad (1.73)$$

при некотором $s=1, 2, \dots, r$. Эта система обладает единственным решением.

Обозначим

$$p_j^{(s)} = \begin{cases} p_j, & j \in C_s; \\ 0, & j \in \bar{C}_s, \end{cases} \quad (1.74)$$

где $\{p_s\}$ — решение выписанной системы при данном s .

Справедливо следующее свойство. При условии, если в некоторый момент t_0 $v(t_0) \in C_s$, то

$$\lim_{t \rightarrow \infty} P(v(t) = j) = p_j^{(s)}; \quad j \in \Omega. \quad (1.75)$$

Если i, j — любые состояния процесса, то можно записать

$$\lim_{t \rightarrow \infty} P(v(t) = j \mid v(0) = i) = \sum_{s=1}^r q_{is} p_j^{(s)}, \quad (1.76)$$

где q_{is} — вероятность попадания процесса в некоторое состояние из множества C_s при условии, что $v(0) = i$.

Вероятности q_{is} находятся из следующей системы уравнений:

$$\lambda_i q_{is} = \sum_{k \in C_s} \lambda_{ik} q_{ks}; \quad i \in C_0, \quad (1.77)$$

где C_0 — множество несущественных состояний процесса $v(t)$.

Для $i \in C_s$ ($1 \leq s \leq r$) имеем:

$$q_{is} = 1; \quad q_{im} = 0; \quad m \neq s. \quad (1.78)$$

Процессы размножения и гибели

Пусть $X = \{0, 1, 2, \dots\}$, причем из любого состояния возможен переход лишь в соседнее состояние. Таким образом

$$\lambda_{ij} = \begin{cases} \lambda_i & \text{при } j = i + 1; \\ \mu_i & \text{при } j = i - 1 \geq 0; \\ 0 & \text{при } j \geq i + 2; \\ 0 & \text{при } j \leq i - 2; \\ 0 & \text{при } j < 0. \end{cases} \quad (1.79)$$

Такой однородный марковский процесс носит название процесса «размножения и гибели»¹.

Процессы размножения и гибели выступают в качестве моделей функционирования многих систем массового обслуживания.

Обозначим

$$q_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \quad (n \geq 1); \quad q_0 = 1. \quad (1.80)$$

Если

$$\sum_{n=0}^{\infty} q_n < \infty; \quad \sum_{n=0}^{\infty} \frac{1}{\lambda_n q_n} = \infty, \quad (1.81)$$

то процесс размножения и гибели обладает эргодическим распределением

$$\lim_{t \rightarrow \infty} P(v(t) = j) = \frac{q_j}{\sum_{n=0}^{\infty} q_n} \quad (j = 0, 1, 2, \dots). \quad (1.82)$$

1.3. ПОЛУМАРКОВСКИЕ ПРОЦЕССЫ. ЛИНЕЙЧАТЫЕ ПРОЦЕССЫ

Марковские процессы с конечным или счетным множеством состояний представляют особый интерес для решения прикладных задач теории массового обслуживания. Однако при этом появление новых заявок и окончание обслуживания заявок, уже имеющих в системе, не должно зависеть от предшествующей истории. Иными словами, должен выполняться марковский принцип независимости будущего от прошедшего при известном настоящем.

Нет нужды доказывать, что в большинстве ситуаций, встречающихся в практике проектирования и обслуживания систем управления, такого рода предпосылки далеки от действительности. В связи с этим приходится использовать случайные процессы более сложного характера.

Обозначим через $v(t)$ некоторый процесс, в любой момент времени описывающий состояние СМО (предпо-

¹ Название заимствовано из приложения к биологии, где подобными процессами описывается динамика различных популяций, например, бактерий.

лагается, что по реализациям случайной функции $v(t)$ можно проследить за всеми изменениями, которые происходят в системе: моментами появления заявок, моментами окончания обслуживания и т. п.). Примером процесса $v(t)$ может служить число заявок, находящихся в СМО в произвольный момент времени t .

Иногда естественно рассматривать процесс $v(t)$ как совокупность нескольких параметров, имеющих тот или иной физический смысл. Так, например, при рассмотрении системы управления с приборами, которые могут отказывать, имеет смысл считать, что процесс $v(t)$ двумерный:

$$v(t) = \{v_1(t), v_2(t)\}, \quad (1.83)$$

где $v_1(t)$ — число заявок в системе в момент t ; $v_2(t)$ — число неисправных приборов в тот же момент времени.

Если известны вероятностные законы, управляющие входящим потоком требований, а также распределение длительности обслуживания и порядок обслуживания заявок, то $v(t)$ задан как случайный процесс.

Перейдем к изложению смысла метода *вложенных цепей Маркова*, предложенного Д. Кендаллом [14, 71].

Изучается некоторый случайный процесс $v(t)$, о котором известно, что он не обладает марковским свойством. Выбираются такие моменты времени $\{t_n\}$, ($t_n < t_{n+1}$), что значения процесса $\{v(t_n)\}$ образуют цепь Маркова. Затем методами, обычными для цепей Маркова, исследуется распределение случайных величин $v(t_n)$. Наконец, по этому распределению делают выводы о свойствах исходного процесса $v(t)$. Во многих случаях этот последний этап исключается, поскольку сами величины $v(t_n)$ дают исчерпывающую информацию о функционировании СМО.

Чаще всего рассматривается случай, когда множество возможных значений процесса $v(t)$, а следовательно, и вложенной цепи Маркова конечно или счетно. Однако можно рассматривать и непрерывное множество состояний. С этой точки зрения метод вложенных цепей Маркова охватывает и теорию случайных блужданий, также используемую в задачах массового обслуживания.

Таким образом, вложенная цепь Маркова — это последовательность значений процесса в специально выбранные моменты времени $\{t_n\}$; эти значения образуют цепь Маркова. Подчеркнем, что моменты $\{t_n\}$, как правило, оказываются не детерминированными, а случай-

ными, зависящими от поведения самого процесса $v(t)$. О моментах $\{t_n\}$ здесь ничего не говорится.

Вложенная цепь Маркова может быть определена и в том случае, когда моменты $\{t_n\}$ образуют сложную статистическую связь, которая не может быть описана цепью Маркова. Важный частный случай выделяется посредством определения *полумарковского процесса*, которое мы приведем, следуя работам [14, 42].

Пусть N —некоторое конечное или счетное множество и пусть для каждого i из этого множества определены: вероятности перехода P_{ij} со свойствами

$$0 \leq P_{ij} \leq 1; \sum_{j \in N} P_{ij} = 1; i \in N; \quad (1.84)$$

распределение вероятностей $F_i(x)$ положительной случайной величины $\xi^{(i)}$

$$F_i(x) = P\{\xi^{(i)} < x\}; F_i(+0) = 0; i \in N. \quad (1.85)$$

Будем считать также заданным распределение случайного вектора (ξ_0, v_0) , первая компонента которого—неотрицательная случайная величина, а вторая—может принимать значения из множества N .

Полумарковский процесс конструируется из введенных объектов по следующему принципу.

Пусть случайный вектор (ξ_0, v_0) принял некоторое значение (x, i) в соответствии с распределением, о котором говорилось выше. Тогда определим $v(t)$ в интервале времени от 0 до ξ_0 как v_0 :

$$v(t) = i; 0 \leq t < x. \quad (1.86)$$

Затем из случайной совокупности N выбирается значение v_1 ; вероятность того, что будет выбрано именно $v_1 = j$ равна P_{ij} (напомним, что i —значение, принятое v_0). После этого выбирается значение случайной величины $\xi_1 = y$ в соответствии с законом распределения $F_j(x)$, что позволяет определить процесс $v(t)$ в более широком промежутке времени:

$$v(t) = j; x \leq t < x + y. \quad (1.87)$$

Вслед за этим производится новый выбор элемента v_2 из множества N . На этот раз вероятность того, что v_2 примет значение k , равна P_{jk} , $k \in N$. Как скоро стало известно, что $v_2 = k$, выбирается значение случайной ве-

личины $\xi_2 = z$ в соответствии с распределением $F_h(x)$, что дает:

$$v(t) = k; \quad x + y \leq t < x + y + z \quad (1.88)$$

и т. д.

В результате такого построения случайный процесс $v(t)$ определится в интервале времени $(0, t^*)$, где

$$t^* = x + y + z + \dots \quad (1.89)$$

Чаще всего $t^* = \infty$, так что процесс определяется при всех $t > 0$. Это и будет полумарковский процесс¹.

Наконец, дадим определение *линейчатого процесса*, которым мы также будем пользоваться в дальнейшем.

Пусть имеется полумарковский процесс $v(t)$, причем с вероятностью единица он продолжается неограниченно долго, т. е. согласно приведенной конструкции $v(t)$ определяется при всех $t > 0$. Тогда для любого $t > 0$ можно определить случайную функцию $\gamma^*(t)$ как время, прошедшее с момента последнего выбора случайной величины. Более точно

$$\gamma^*(t) = t - (\xi_0 + \xi_1 + \dots + \xi_n), \quad (1.90)$$

где n находится из условия

$$\xi_0 + \xi_1 + \dots + \xi_n < t \leq \xi_0 + \xi_1 + \dots + \xi_n + \xi_{n+1}. \quad (1.91)$$

Чтобы это определение было справедливо всегда, положим, что при $t < \xi_0$ $\gamma^*(t) = t$.

Легко видеть, что двумерный случайный процесс $\{v(t), \gamma^*(t)\}$ будет марковским. Действительно, пусть известно, что при фиксированном t $v(t) = i$, $\gamma^*(t) = x$. Это означает, в частности, что последний «цикл» длится уже x единиц времени. Но тогда вероятность окончания цикла в интервале $(t, t + y)$ составляет

$$\frac{F_i(x + y) - F_i(x)}{1 - F_i(x)}.$$

До окончания цикла процесс $v(t)$ не изменит своего значения, а процесс γ^* будет все время возрастать с еди-

¹ Полумарковский процесс определяется и несколько иначе: считают, что время пребывания в данном состоянии зависит от последующего состояния процесса. Оба определения эквивалентны. Для перехода от одного из них к другому может потребоваться лишь изменение множества состояний процесса.

ничной скоростью. Если же цикл окончился, то известно, что может произойти дальше: с вероятностью P_{ij} процесс $v(t)$ переходит к значению $v(t) = j$ и все дальнейшее определяется независимо от того, что было до момента t .

Сделаем еще одно замечание. Если для некоторых i из множества N функция распределения $F_i(x)$ имеет вид

$$F_i(x) = 1 - e^{-\lambda_i x}; \quad x \geq 0,$$

то при $v(t) = i$ не нужно определять вторую компоненту процесса, т. е. $\gamma^*(t)$: остаток длительности цикла, как легко видеть, зная основное свойство показательного распределения, не зависит от $\gamma^*(t)$.

Таким образом можно рассматривать процесс $\xi^*(t)$, в который обязательно включается компонента $v(t)$; вторая компонента, именно $\gamma^*(t)$, определяется лишь для тех случаев, когда $F_i(x)$ имеет вид, отличный от показательного [14].

Термин «линейчатый марковский процесс» для процессов несколько иного типа был предложен Ю. К. Беляевым [8]. Заметим также, что линейчатый процесс — есть функционал от полумарковского процесса.

1.4. ПРОЦЕССЫ ВОССТАНОВЛЕНИЯ

Пусть $\xi_0, \xi_1, \dots, \xi_n, \dots$ — последовательность неотрицательных независимых случайных величин, где ξ_1, ξ_2, \dots обладают общей функцией распределения $F(x)$, а ξ_0 — функцией распределения $F_0(x)$. Последовательность моментов времени $t_n = \xi_0 + \dots + \xi_n$ ($n = 0, 1, 2, \dots$) называется *процессом восстановления*. В частном случае, когда $F_0(x) = F(x)$, процесс восстановления называется *простым*.

Если некоторая из случайных величин t_n равна t , говорят, что в момент t происходит восстановление. С процессом восстановления связывается случайный процесс N_t , определяемый как число восстановлений в интервале $(0, t)$. Если $0 < t_0 < t_1 < \dots < t_n < \dots$, тогда $N_t = 0$ при $0 < t < t_0$; $N_t = 1$ при $t_0 < t < t_1, \dots$; $N_t = n$ при $t_{n-1} < t < t_n$ и т. д. В этом случае траектория процесса восстановления представляет собой ступенчатую функцию, возрастающую единичными скачками. Если некоторые моменты восстановления совпадают (скажем $t_1 = t_2 = t_3$), возможны и большие скачки.

Основой изучения процессов восстановления является понятие *функции восстановления*.

Функцией восстановления данного процесса восстановления называется функция

$$H(t) = M[N_t], \quad t \geq 0. \quad (1.92)$$

Выведем формулу для функции восстановления. Пусть $t \geq 0$ — фиксированное число и $U(x) = 1$ при $x < t$, $U(x) = 0$ при $x \geq t$. Тогда, если $U(t_n) = 1$, значит n -е восстановление произошло ранее t . Следовательно, $\sum_{n=0}^{\infty} U(t_n)$ представляет собой общее число восстановлений, происшедших ранее t , или, что то же, N_t .

Таким образом

$$H(t) = M \left[\sum_{n=0}^{\infty} U(t_n) \right] = \sum_{n=0}^{\infty} M[U(t_n)]. \quad (1.93)$$

Однако, $U(t_n)$ — случайная величина, принимающая одно из двух возможных значений: «0» или «1». Поэтому

$$M[U(t_n)] = P(t_n < t) = P(\xi_0 + \xi_1 + \dots + \xi_n < t). \quad (1.94)$$

Окончательно

$$H(t) = \sum_{n=0}^{\infty} P(\xi_0 + \xi_1 + \dots + \xi_n < t). \quad (1.95)$$

Перейдем к преобразованиям Лапласа-Стилтьеса. Если $F(t)$ — неубывающая функция, то преобразованием Лапласа-Стилтьеса этой функции называется функция комплексного переменного $\varphi(s) = \int_0^{\infty} e^{-st} dF(t)$; если $F(\infty) < \infty$, тогда $\varphi(s)$ определена при всех s , для которых $\operatorname{Re} s \geq 0$.

Дадим также следующее объяснение рассматриваемому преобразованию. Если функция $F(t)$ имеет скачки величины Δ_n в точках x_n ($0 < x_1 < x_2 < \dots$), а в интервалах между этими точками обладает непрерывной производной $F'(t)$, тогда

$$\begin{aligned} \varphi(s) = & \sum_n e^{-sx_n} \Delta_n + \int_0^{x_1} e^{-st} F'(t) dt + \\ & + \int_{x_1}^{x_2} e^{-st} F'(t) dt + \dots \end{aligned} \quad (1.96)$$

Обозначим

$$\left. \begin{aligned} \varphi_0(s) &= \int_0^{\infty} e^{-st} dF_0(t); & \varphi(s) &= \int_0^{\infty} e^{-st} dF(t); \\ \psi(s) &= \int_0^{\infty} e^{-st} dH(t). \end{aligned} \right\} \quad (1.97)$$

Введем также функции распределения

$$F_n(t) = P(\xi_0 + \xi_1 + \dots + \xi_n < t) \quad (1.98)$$

и их преобразования Лапласа-Стилтьеса

$$\varphi_n(s) = \int_0^{\infty} e^{-st} dF_n(t). \quad (1.99)$$

Тогда символически можно записать

$$dH(t) = \sum_{n=0}^{\infty} dF_n(t), \quad (1.100)$$

откуда

$$\psi(s) = \sum_{n=0}^{\infty} \varphi_n(s). \quad (1.101)$$

Известно, что преобразование Лапласа-Стилтьеса суммы независимых случайных величин равно произведению преобразований слагаемых. Тогда

$$\varphi_n(s) = \varphi_0(s) \varphi^n(s) \quad (1.102)$$

и

$$\psi(s) = \varphi_0(s) \sum_{n=0}^{\infty} \varphi^n(s) = \frac{\varphi_0(s)}{1 - \varphi(s)}. \quad (1.103)$$

Пусть имеется простой процесс восстановления. Соответствующую ему функцию восстановления обозначим $H_0(t)$. Согласно предыдущему

$$H_0(t) = M \left[\sum_{n=0}^{\infty} U(t_n) \right] = M[U(t_0)] + M \left[\sum_{n=1}^{\infty} U(t_n) \right]. \quad (1.104)$$

Поскольку $M[U(t_0)] = F(t)$, можно записать

$$H_0(t) = F(t) + \int_0^{\infty} M \left[\sum_{n=1}^{\infty} U(t_n) | t_0 = x \right] dF(x). \quad (1.105)$$

Если $t_0 \geq t$, то и все $t_n \geq t$, так что $U(t_n) = 0$. Таким образом, в последнем интеграле можно взять в качестве верхнего предела t . Далее, если известно, что $t_0 = x$, случайная величина t_n имеет то же распределение, что и случайная величина $t_{n-1} + x$.

Итак,

$$\begin{aligned} M \left[\sum_{n=1}^{\infty} U(t_n) | t_0 = x \right] &= M \left[\sum_{n=1}^{\infty} U(t_{n-1} + x) \right] = \\ &= \sum_{n=0}^{\infty} P(t_n < t - x) = H_0(t - x). \end{aligned} \quad (1.106)$$

Окончательно

$$H_0(t) = F(t) + \int_0^t H_0(t-x) dF(x).$$

Найденное соотношение носит название *уравнения восстановления*.

Возвратимся теперь к исходному процессу восстановления, для которого функция $F_0(t)$ может быть отличной от $F(t)$. Можно записать

$$H(t) = F_0(t) + \int_0^t H_0(t-x) dF_0(x). \quad (1.107)$$

Уравнение для вероятности события, связанного с процессом восстановления

Пусть нас интересует вероятность p некоторого события A , определенного следующим образом. Событие A может произойти лишь в один из моментов t_n ($n \geq 0$). Если событие A не произошло в моменты t_i ($i < n$), то в момент t_n оно произойдет с вероятностью $a(t)$, если известно, что $t_n = t$ независимо от значений t_i ($i < n$).

Пусть, например, t_n — моменты отказа некоторого устройства, функционирующего на отрезке времени $[0, T]$

в переменных условиях. Если отказ произошел в момент t , то этот отказ приведет к катастрофе с вероятностью $\alpha(t)$. (Естественно положить $\alpha(t) = 0$ при $t > T$). Тогда катастрофа как раз и будет событием с описанными свойствами.

Для простоты вывода предположим, что ξ_n ($n \geq 0$) обладают соответствующими плотностями. Пусть $q_n(t) dt$ обозначает вероятность события, состоящего в том, что $t < t_n < t + dt$ и в моменты t_0, \dots, t_n событие A не произошло. Тогда

$$q = 1 - p = \lim_{n \rightarrow \infty} \int_0^{\infty} q_n(x) dx. \quad (1.108)$$

Далее имеем интегральное соотношение

$$q_n(t) = \beta(t) \int_0^t q_{n-1}(x) F'(t-x) dx, \quad n \geq 1, \quad (1.109)$$

где $\beta(t) = 1 - \alpha(t)$.

Далее можно записать

$$p = \int_0^{\infty} \alpha(t) F_0'(t) dt + \sum_{n=0}^{\infty} \int_0^{\infty} \int_0^{\infty} q_n(t) F'(x) \alpha(t+x) dt dx. \quad (1.110)$$

Первый член правой части соответствует случаю, когда событие A происходит в момент t_0 ; выражение $q_n(t) F'(x) \alpha(t+x) dt dx$ есть вероятность следующего события: n -е восстановление произошло в момент t_n , где $t < t_n < t + dt$; событие A не произошло в моменты t_0, t_1, \dots, t_n ; $(n+1)$ -е восстановление и событие A произошли в момент t_{n+1} , где $t_n + x < t_{n+1} < t_n + x + dx$.

Поменяв местами знаки суммы и интеграла, найдем

$$p = \int_0^{\infty} \alpha(t) F_0'(t) dt + \int_0^{\infty} \int_0^{\infty} \left[\sum_{n=0}^{\infty} q_n(t) \right] F'(x) \alpha(t+x) dt dx. \quad (1.111)$$

Обозначим

$$q(t) = \sum_{n=0}^{\infty} q_n(t).$$

Далее имеем

$$q_0(t) = F_0'(t) \beta(t).$$

Тогда

$$q(t) = F_0'(t) \beta(t) + \beta(t) \int_0^t \sum_{n=1}^{\infty} q_{n-1}(x) F'(t-x) dx \quad (1.112)$$

или

$$q(t) = \left[F_0'(t) + \int_0^t q(x) F'(t-x) dx \right] \beta(t). \quad (1.113)$$

Определив функцию $q(t)$ согласно интегральному уравнению (1.113), вычислим p по формуле

$$p = \int_0^{\infty} \alpha(t) F_0'(t) dt + \int_0^{\infty} \int_0^{\infty} q(t) F'(x) \alpha(t+x) dt dx. \quad (1.114)$$

Таким образом, поставленная задача решена.

Для последующих исследований важны следующие теоремы.

1. Элементарная теорема восстановления.

Пусть $\tau = M[\xi_n] < \infty$ при $n \geq 1$. Тогда

$$\lim_{t \rightarrow \infty} \frac{H(t)}{t} = \frac{1}{\tau}. \quad (1.115)$$

2. Узловая теорема восстановления.

Пусть $Q(t)$ — невозрастающая ограниченная функция, для которой $\int_0^{\infty} Q(t) dt < \infty$. Тогда, если $F(t)$ — нерешетчатая функция распределения¹, то

$$\lim_{t \rightarrow \infty} \int_0^t Q(t-x) dH(x) = \frac{1}{\tau} \int_0^{\infty} Q(t) dt. \quad (1.116)$$

Более подробно эти вопросы рассмотрены в книге [22].

¹ Имеется в виду, что либо ξ_n — не дискретная, либо ξ_n — дискретная случайная величина, но не все ее значения соизмеримы.

Стационарные процессы восстановления

Предположим, что для любого фиксированного $x > 0$ распределения числа восстановлений в интервале $(t, t+x)$ не зависит от t ($t \geq 0$). Процесс восстановления с таким свойством называется *стационарным процессом восстановления*. Если задана $F(t) = P(\xi_n < t)$ при $n \geq 1$, причем $\tau = M\xi_1 < \infty$, то необходимым и достаточным условием стационарности процесса восстановления является существование у случайной величины ξ_0 плотности вида

$$F_0'(t) = \frac{1}{\tau} [1 - F(t)]; \quad t > 0. \quad (1.117)$$

Если же $\tau = \infty$, процесс восстановления не может быть стационарным.

Для стационарных процессов восстановления $H(t) = \frac{t}{\tau}$ ($t \geq 0$). Если $F(+0) = 0$, то вероятность восстановления в интервале $(t, t+\Delta t)$ равна $\frac{1}{\tau} \Delta t + o(\Delta t)$; вероятность более одного восстановления в интервале $(t, t+\Delta t)$ составляет $o(\Delta t)$.

Обозначим через γ_t время от момента t до следующего после этого момента восстановления.

Имеем

$$t + \gamma_t = t_{N_t+1}. \quad (1.118)$$

Для стационарного процесса восстановления с $F(+0) = 0$ случайная величина γ_t имеет распределение, не зависящее от t при $t \geq 0$ и обладающее плотностью

$$p_{\gamma_t}(x) = \frac{1}{\tau} [1 - F(x)]; \quad x > 0. \quad (1.119)$$

Стационарный процесс восстановления с $F(+0) = 0$ иначе называется потоком однородных событий типа Пальма. Теория таких потоков развита А. Я. Хинчиным [46].

Случайная замена времени в процессах восстановления

Пусть $\{t_n\}$ — произвольный процесс восстановления, а $x(t)$ — неубывающий, не зависящий от $\{t_n\}$ случайный процесс, для которого с вероятностью равной единице

$x(0) \geq 0$. Дополним график функции $x=x(t)$ (рис. 1.2) в точках разрыва отрезками прямых, параллельных оси Ox так, чтобы в результате получилась непрерывная кривая на плоскости (t, x) . Проведем в этой плоскости пря-

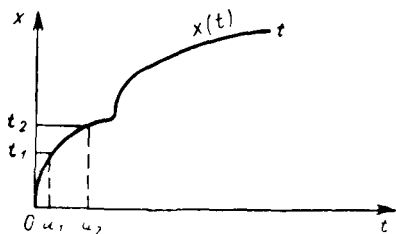


Рис. 1.2. График случайной функции $x(t)$

мую $x=t_n$. Наименьшее значение t , при котором $x(t)=t_n$, обозначим u_n . Множество моментов времени $\{u_n\}$ назовем процессом восстановления с измененным временем. Если, в частности, $x(t) \equiv t_1$, тогда $u_n \equiv t_1, n \geq 0$.

Вначале дадим физическую интерпретацию замены времени.

Представим себе, что имеется бесконечный запас элементов с ресурсами надежности $\xi_0, \xi_1, \dots, \xi_n, \dots$. Когда ресурс надежности n -го элемента исчерпывается, этот элемент отказывает и немедленно заменяется $(n+1)$ -м элементом. Обозначим через $\gamma(t)$ ресурс надежности элемента, служащего в момент t . Если ресурс надежности исчерпывается со скоростью, равной единице, то $\gamma(t)$ будет функцией, которая в интервалах (t_n, t_{n+1}) убывает с единичной скоростью (рис. 1.3, а). В этом случае $\gamma(t) = \gamma_t$ (см. выше). Если же ресурс надежности исчерпывается со скоростью $y(t)$, тогда $\gamma'(t) = -y(t)$ в интервалах между скачками процесса $\gamma(t)$. Если процесс $x(t)$ имеет производную, то $x'(t) = y(t)$.

Так, например, представим себе устройство периодического действия, отказы которого возможны в интервалах длины τ_1 , сменяющихся интервалами длины τ_0 , в которых отказы невозможны. Поведение функции $\gamma(t)$ в этом случае будет таким, как показано на рис. 1.3, б. Описанная ситуация характерна для практики эксплуатации сложных систем.

Предположим, что процесс восстановления является стационарным и $F(+0) = 0$. Пусть, далее $x(t)$ непрерывно дифференцируема с вероятностью, равной единице. Обозначим через $\lambda(t)dt$ вероятность того, что в интервале $(t, t+dt)$ произойдет восстановление для процесса $\{u_n\}$. Тогда

$$\lambda(t) = \frac{1}{\tau} M[x'(t)]. \quad (1.120)$$

Фиксируем некоторые $T_1 < T_2$ и обозначим через $q(T_1, T_2)$ вероятность того, что в интервале (T_1, T_2) не произойдет ни одного восстановления процесса $\{u_n\}$.

Для соответствующей формулы понадобятся функции:

$$\Phi(z) = \frac{1}{\tau} \int_0^{\infty} [1 - F(x)] dx; \quad (1.121)$$

$$G(z) = P\{(x(T_2) - x(T_1)) < z\}. \quad (1.122)$$

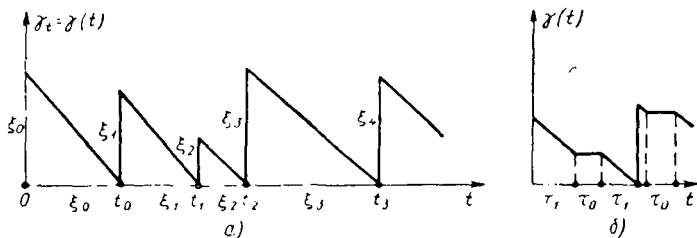


Рис. 1.3. Ресурс надежности элемента как функция времени

Формула для $q(T_1, T_2)$ имеет вид

$$q(T_1, T_2) = \int_0^{\infty} \Phi(z) dG(z). \quad (1.123)$$

Рассмотрим произвольное число моментов времени $T_1 < T_2 < \dots < T_m$ и обозначим через $\lambda(T_1, \dots, T_m) dT_1 \dots dT_m$ вероятность события, состоящего в том, что в каждом из интервалов $(T_1, T_1 + dT_1), \dots, (T_m, T_m + dT_m)$ происходит одно восстановление процесса $\{u_n\}$.

Обозначим через $p(a_1, \dots, a_{m-1}; b_1, \dots, b_m)$ совместную плотность вероятности случайных величин

$$x(t_2) - x(t_1), \quad x(t_3) - x(t_2), \quad \dots, \quad x(t_m) - x(t_{m-1}), \\ x'(t_1), \quad \dots, \quad x'(t_m),$$

предположив, что такая плотность существует. При этом предположении справедлива формула

$$\lambda(T_1, \dots, T_m) = \\ = \int \dots \int b_1 \dots b_m H(a_1) \dots H'(a_{m-1}) p(a_1, \dots, a_{m-1}; \\ b_1, \dots, b_m) da_1 \dots db_m, \quad (1.124)$$

где $H(t)$ — функция восстановления процесса $\{t_n\}$.

Пусть при тех же условиях $\mu(T_1, \dots, T_m) dT_1 \dots dT_m$ — вероятность события, состоящего в том, что в каждом из интервалов $(T_1, T_1 + dT_1); \dots; (T_m, T_m + dT_m)$ произойдет по одному восстановлению процесса $\{u_n\}$, а в промежуточных интервалах $(T_1 + dT_1, T_2); \dots; (T_{m-1}, T_{m-1} + dT_{m-1}, T_m)$ не будет ни одного восстановления. Тогда справедлива формула (в предположении, что ξ_n обладают плотностью)

$$\begin{aligned} & \mu(T_1, \dots, T_m) = \\ & = \frac{1}{\tau} \int \dots \int b_1 b_2 \dots b_m F'(a_1) \dots F'(a_{m-1}) p(a_1, \dots, a_{m-1}; \\ & \quad b_1, \dots, b_m) da_1 \dots db_m. \end{aligned} \quad (1.125)$$

Альтернирующие процессы восстановления

Пусть $v(t)$ — определенный при $t \geq 0$ случайный процесс, траектории которого представляют собой ступенчатые функции со значениями «0» и «1». Допустим, что существует последовательность моментов $t_1 < t_2 < \dots < t_n < \dots$, в каждый из которых процесс изменяет свое состояние.

Возможны два случая: 1) $v(t) = 0, 0 < t < t_1$; 2) $v(t) = 1, 0 < t < t_1$. В первом случае $v(t) = 0$ в интервалах $(t_2, t_3), (t_4, t_5), \dots$; $v(t) = 1$ в интервалах $(t_1, t_2), (t_3, t_4), \dots$. Во втором случае $v(t) = 0$ в интервалах $(t_1, t_2), (t_3, t_4), \dots$; $v(t) = 1$ в интервалах $(t_2, t_3), (t_4, t_5), \dots$.

В обоих случаях предположим, что t_1 и $t_2 - t_1, \dots$ — независимые случайные величины. Если в интервале $(0, t_1)$ $v(t) = 1$ с вероятностью p , то $P(t_1 < t) = G_0(t)$; если $v(t) = 0$ с вероятностью $q = 1 - p$, то $P(t_1 < t) = F_0(t)$.

Длина каждого интервала (t_n, t_{n+1}) ($n \geq 1$), если в этом интервале $v(t) = 0$ распределена по закону $P(t_{n+1} - t_n < t) = F(t)$; если же $v(t) = 1$, то закон распределения $P(t_{n+1} - t_n < t) = G(t)$. Определенный таким образом процесс $v(t)$ называется *альтернирующим процессом восстановления*.

Обозначим через $p(t)$ вероятность того, что $v(t) = 1$. Для того, чтобы вычислить эту вероятность, введем вспомогательный альтернирующий процесс восстановления, у которого $F(t)$ и $G(t)$ — те же, что и у процесса $v(t)$, но в качестве p и $G_0(t)$ взяты соответственно «1» и $G(t)$. Обозначим через $p_0(t)$ вероятность того, что вспомогательный процесс принимает значение «1» в момент t .

Пусть $H_0(t)$ — функция восстановления простого процесса восстановления с функцией распределения ξ_n вида

$$P(\xi_n < t) = \int_0^t F(t-x) dG(x). \quad (1.126)$$

Тогда

$$p_0(t) = 1 - G(t) + \int_0^t [1 - G(t-x)] dH_0(x). \quad (1.127)$$

Функцию $p(t)$ можно найти по формуле

$$p(t) = p[1 - G_0(t)] + \int_0^t p_0(t-x) dR(x), \quad (1.128)$$

где

$$R(t) = p \int_0^t G_0(t-x) dF(x) + qF_0(t).$$

Если хотя бы одно из распределений $F(t)$ или $G(t)$ является перешетчатым, то при $t \rightarrow \infty$

$$\lim p(t) = \frac{\tau_1}{\tau_0 + \tau_1}, \quad (1.129)$$

где $\tau_0 = \int_0^{\infty} t dF(t)$, $\tau_1 = \int_0^{\infty} t dG(t)$ и хотя бы одна из величин τ_0 или τ_1 конечна.

1.5. ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ПРОЦЕССОМ ВОССТАНОВЛЕНИЯ

Пусть имеется некоторое множество S управлений s . Если мы выберем управление $s \in S$, то происходит некоторый процесс случайной длительности $\xi(s)$ с ограниченным по всему множеству S положительным математическим ожиданием $\tau(s)$ и в течение этого времени вырабатываются случайные значения $\eta_1(s), \dots, \eta_m(s)$ некоторых числовых характеристик.

Так, например, допустим, что поставлена в эксплуатацию новая система. Управление s определяется режимом ее эксплуатации, профилактическими работами, функци-

Записанные равенства можно интерпретировать следующим образом. Если на отрезке времени $[0, T]$ уложились некоторое число циклов, то суммарное значение характеристики η_j за время T есть сумма значений этой характеристики по данным циклам плюс некоторая часть значений этой характеристики по циклу, не уложившемуся полностью на отрезке $[0, T]$.

Теперь сформулируем следующую задачу. Обозначим

$$\frac{1}{T} \eta_j(T) = z_j(u, T); \quad 1 \leq j \leq m, \quad (1.131)$$

где $u = (s_0, s_1, \dots, s_n, \dots)$.

Обозначим через S_0 множество стратегий u , для которых существуют такие постоянные $z_j(u)$, что для любого $\varepsilon > 0$

$$P(|z_j(u, T) - z_j(u)| < \varepsilon) \rightarrow 1 \quad (T \rightarrow \infty).$$

Требуется выбрать стратегию $u \in S_0$ таким образом, чтобы вектор $\{z_1(u), \dots, z_{m-1}(u)\}$ принадлежал некоторому выпуклому множеству Φ и при этом функция $z_m(u)$ была бы минимально возможной.

Типичным примером является требование минимизации $z_m(u)$ при условии, что

$$a_j \leq z_j(u) \leq b_j; \quad 1 \leq j \leq m-1.$$

Будем считать, что множество S замкнуто и допускает операцию смешивания. Последнее означает следующее. Пусть s_1, \dots, s_N — управления из множества S . Тогда существует рандомизированное управление, при котором s_i применяется с вероятностью p_i , где p_1, \dots, p_N — любое распределение вероятностей. Нам понадобится лишь одно следствие этого свойства, а именно такое: множество векторов $\{\tau(s), \psi_1(s), \dots, \psi_m(s)\}$ является выпуклым, где

$$M[\eta_j(s)] = \psi_j(s); \quad s \in S; \quad 1 \leq j \leq m. \quad (1.132)$$

Будем считать, что эти величины конечны для любого $s \in S$ и пусть

$$\psi(s) = (\psi_1(s), \dots, \psi_m(s)). \quad (1.133)$$

Далее можно определить

$$\psi_j^*(s) = \frac{\psi_j(s)}{\tau(s)}; \quad 1 \leq j \leq m.$$

Предположим, что выполнены следующие условия.

1. Для любой последовательности управлений $\{s_n\}$ последовательность случайных векторов $\{(\xi(s_i), \eta_1(s_i), \dots, \eta_m(s_i))\}$ удовлетворяет закону больших чисел в следующем виде.

Если $\sum_{i=0}^n \tau(s_i) \rightarrow \infty$, то

$$P\left(\left|\sum_{i=0}^n \xi(s_i) - M \sum_{i=0}^n \xi(s_i)\right| > \varepsilon \sum_{i=0}^n \tau(s_i)\right) \rightarrow 0 \quad (n \rightarrow \infty);$$

$$P\left(\left|\sum_{i=0}^n \eta_j(s_i) - M \sum_{i=0}^n \eta_j(s_i)\right| > \varepsilon \sum_{i=0}^n \tau(s_i)\right) \rightarrow 0 \quad \text{при лю-}$$

бом $\varepsilon > 0$ ($n \rightarrow \infty$).

2. $M[|\eta_j(s)|] \leq c\tau(s)$; $s \in S$,

где $c > 0$ — константа.

Предположим, что

$$\psi_m^*(s^*) = \min_{s \in H} \frac{\psi_m(s)}{\tau(s)} = \lim_{\varepsilon \rightarrow 0} \min_{s \in H_\varepsilon} \frac{\psi_m(s)}{\tau(s)}. \quad (1.133)$$

где s^* — некоторый элемент множества S ;

H — множество тех s , для которых $(\psi_1^*(s), \dots, \psi_{m-1}^*(s)) \in \Phi$;
 H_ε — множество тех s , для которых $(\psi_1^*(s), \dots, \psi_{m-1}^*(s))$ принадлежит ε -окрестности Φ_ε множества Φ^1 .

Теорема 1. Стратегия $u^* = (s^*, s^*, \dots, s^*, \dots)$ принадлежит S_0 и является решением поставленной задачи.

Эта теорема носит экзистенциальный характер. Следующая же теорема указывает способ построения оптимальной стратегии.

Теорема 2. В условиях теоремы 1 существуют такие вещественные числа $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$, что задача минимизации $\psi_m^*(s)$ при условии $(\psi_1^*(s), \dots, \psi_{m-1}^*(s)) \in \Phi$ эквивалентна задаче минимизации суммы

$$\Lambda(s) = \lambda_0 \tau(s) + \sum_{j=1}^m \lambda_j \psi_j(s)$$

по множеству S .

¹ Совокупность векторов, каждый из которых либо входит в Φ , либо находится на расстоянии, меньшем ε , от некоторого вектора из множества Φ , называется ε -окрестностью множества Φ .

Приведенная теорема позволяет во многих практических случаях сводить задачу построения оптимальной решающей функции с неаддитивной функцией убытка к задаче с аддитивной функцией убытка.

В качестве примера рассмотрим задачу выбора оптимального плана профилактики, решенную в гл. IV иным методом.

Пусть время X жизни элемента распределено по закону $F(t) = P(X < t)$. Если элемент отказал, то он заменяется новым, причем среднее время замены равно T_s . Если произведена профилактика ранее отказа элемента, то средняя длительность профилактики равна $T_m < T_s$. Требуется оптимальным образом выбрать функцию распределения $G(x)$ момента Y проведения профилактики с тем, чтобы максимизировать вероятность застать элемент в исправном состоянии.

Обозначим через τ среднюю длительность цикла между возвращениями элемента в исходное состояние, а через ψ — среднее время пребывания элемента в нерабочем состоянии (замены или профилактики) в течение одного цикла. Тогда готовность элемента

$$P = \frac{P(X < Y)T_s + P(X > Y)T_m}{\tau}.$$

По теореме 2 максимизация P равнозначна максимизации функции

$$\begin{aligned} \Lambda &= \lambda_0 \tau + \lambda_1 [P(X < Y)T_s + P(X > Y)T_m] = \\ &= \lambda_0 \left\{ \int_0^{\infty} [1 - F(t)][1 - G(t)] dt + P(X < Y)T_s + P(X > Y)T_m \right\} + \\ &+ \lambda_1 [P(X < Y)T_s + P(X > Y)T_m] = \lambda_0 \int_0^{\infty} [1 - F(t)][1 - G(t)] dt + \\ &+ (\lambda_0 + \lambda_1) \left\{ T_s \int_0^{\infty} [1 - G(t)] dF(t) + T_m \left(1 - \int_0^{\infty} [1 - G(t)] dF(t) \right) \right\} = \\ &= \lambda_0 \int_0^{\infty} [1 - F(t)][1 - G(t)] dt + (\lambda_0 + \lambda_1)(T_s - T_m) \int_0^{\infty} [1 - G(t)] \times \\ &\quad \times dF(t) + (\lambda_0 + \lambda_1) T_m = \int_0^{\infty} \{ \lambda_0 [1 - F(t)] + \\ &\quad + (\lambda_0 + \lambda_1)(T_s - T_m) F'(t) \} [1 - G(t)] dt + (\lambda_0 + \lambda_1) T_m. \end{aligned}$$

Если λ_0 и $\lambda_0 + \lambda_1$ одного знака, то решение будет вырожденным. Так, если $\lambda_0 > 0$ и $\lambda_0 + \lambda_1 > 0$, следует положить $G(t) = 0$ для всех t , т. е. время профилактики равно ∞ .

Если $\lambda_0 < 0$ и $\lambda_0 + \lambda_1 < 0$, нужно положить $G(t) = 1$ при всех $t > 0$. Очевидно, оба этих случая не реальные.

Если $\lambda_0 < 0$ и $\lambda_0 + \lambda_1 > 0$, тогда решение будет одним из двух предыдущих.

Остается случай $\lambda_0 > 0$, а $\lambda_0 + \lambda_1 < 0$.

Если $\frac{F'(t)}{1 - F(t)}$ монотонно возрастающая функция, то уравнение

$$\lambda_0 [1 - F(t)] + (\lambda_0 + \lambda_1)(T_s - T_m) F'(t)$$

имеет единственный корень $\gamma = \gamma\left(\frac{\lambda_1}{\lambda_0}\right)$ в некотором интервале изменения λ_1/λ_0 .

Поскольку при $t < \gamma$ выражение в фигурных скобках положительно, а при $t > \gamma$ отрицательно, то оптимальным будет выбор

$$G(t) = \begin{cases} 0 & \text{при } t \leq \gamma; \\ 1 & \text{при } t > \gamma. \end{cases}$$

Остается рассмотреть первоначальное выражение для P , в которое подставлено последнее выражение $G(t)$, и найти его максимум по γ .

1.6. ПРАВИЛА ОСТАНОВКИ

Предположим, что наблюдается некоторая последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$. *Правилом остановки* процесса наблюдения называется функция $v(\xi_1, \xi_2, \dots, \xi_n, \dots)$, принимающая значения $0, 1, 2, \dots$ и имеющая следующий смысл. Пусть $v(x_1, x_2, \dots, x_n, \dots) = m$. Тогда, если $\xi_1 = x_1, \dots, \xi_m = x_m$, процесс наблюдения случайной последовательности заканчивается на m -м шаге, т. е. наблюдаются только случайные величины ξ_1, \dots, ξ_m , а величины $\xi_{m+1}, \xi_{m+2}, \dots$ уже не наблюдаются.

Чтобы правило остановки имело реальный смысл, нужно потребовать, чтобы $v(x_1, \dots, x_m, \dots)$ не зависела от x_{m+1}, x_{m+2}, \dots при условии, что $v(x_1, \dots, x_m, \dots) = m$. Другими словами, решение об остановке процесса наблюдений должно приниматься на основании уже имеющих, а не будущих наблюдений,

Если $v(\xi_1, \xi_2, \dots, \xi_n, \dots) = 0$, то это означает, что принято решение вообще не производить наблюдений над случайными величинами.

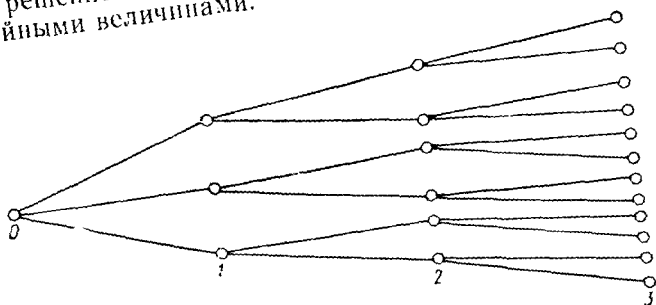


Рис. 1.4. «Дерево» процесса реализации случайных величин

Правила остановки хорошо представить в том случае, когда каждая из случайных величин может принимать лишь конечное число значений. В этом случае, как

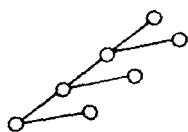


Рис. 1.5. К рассмотрению правила остановки

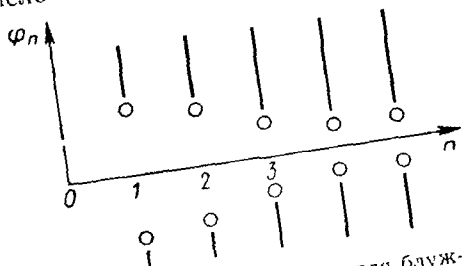


Рис. 1.6. Система барьеров для блуждающей частицы

показано на рис. 1.4, процесс реализации случайных величин ξ_1, ξ_2, \dots можно представить в виде «дерева». Каждой вершине 1-го ранга соответствует некоторое значение ξ_1 ; каждой вершине 2-го ранга — некоторое значение пары (ξ_1, ξ_2) и т. д. Правило остановки в этом случае можно интерпретировать как «обрубывание» веток «дерева» (рис. 1.5).

Нужно заметить, что даже при небольшом числе возможных значений случайных величин ξ_n «дерева» «бы-

стро разрастаются», и задание правила остановки с помощью «дерева» становится весьма громоздким. В целях упрощения можно использовать следующий метод (при этом во многих важных случаях не сужаются возможности построения оптимальных решений). Определяют некоторые функции $\varphi_n = \varphi_n(\xi_1, \dots, \xi_n)$, а, останавливать ли процесс наблюдения на n -м шаге, решают лишь на основании значения φ_n .

В этом случае правило остановки можно представить себе как некоторую систему барьеров для блуждающей частицы (рис. 1.6). Если блуждающая точка (n, φ_n) попадает на один из барьеров при $n = m$, то это означает, что процесс наблюдения заканчивается на m -м шаге.

Понятно, что для любой последовательности наблюдаемых случайных величин можно указать бесчисленное множество правил остановки. Каким же образом сравнивать эти правила и выбирать среди них оптимальные?

Допустим, что для любого $n \geq 0$ и любых x_1, \dots, x_n определена некоторая функция убытка $y = y_n(x_1, \dots, x_n)$. Эта функция интерпретируется как убыток, который будет понесен в случае, если процесс наблюдения остановлен на n -м шаге и при этом первые n наблюдений приняли значения: $\xi_1 = x_1, \dots, \xi_n = x_n$. Предположим, что нам известен вероятностный закон, управляющий образованием случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$. Тогда, коль скоро задано правило остановки $v(\xi_1, \xi_2, \dots)$, в принципе можно вычислить среднюю характеристику

$$I(v) = M [y_{v(\xi_1, \xi_2, \dots)}(\xi_1, \dots, \xi_v)], \quad (1.134)$$

т. е. математическое ожидание значения процесса y_n в момент остановки процесса наблюдений.

Байесовский принцип оптимальности правила остановки состоит в выборе такого правила остановки v^* , для которого

$$I(v^*) = \min_v I(v). \quad (1.135)$$

При этом обычно правила остановки, для которых нельзя придать смысл указанному выше математическому ожиданию, исключаются из рассмотрения.

Правило v^* называется *оптимальным (байесовским) правилом остановки процесса наблюдения*.

Иногда удобно вместо функции убытка задавать так называемую функцию пользы. В этом случае оптимальным правилом остановки будет то, при котором $I(v)$ достигает максимума.

Фундаментальную роль в теории оптимальных правил остановки играет система рекуррентных уравнений для минимального риска. Приведем эти уравнения, дав в общих чертах схему их вывода.

Пусть используется некоторое правило остановки v . Каждой возможной траектории случайной последовательности $(\xi_1, \xi_2, \dots, \xi_n, \dots)$ соответствует некоторое значение убытка $y_v(\xi_1, \dots, \xi_n)$, где $v = v(\xi_1, \xi_2, \dots)$. Выше было отмечено, что оптимальное правило состоит в минимизации $I(v)$, т. е. математического ожидания этого убытка.

Предположим теперь, что получены наблюдения $\xi_1 = x_1, \dots, \xi_n = x_n$. Тогда поведение после n -го наблюдения должно быть таким, чтобы минимизировать условное математическое ожидание $q^{(v)}(x_1, \dots, x_n)$ убытка при условии, что получены указанные наблюдения. Функция $q^{(v)}(x_1, \dots, x_n)$ называется риском после n наблюдений. Очевидно, вид этой функции зависит от правила остановки v . Минимум данной функции по всем допустимым v называется *оптимальным риском* и обозначается $q(x_1, \dots, x_n)$.

Получив n наблюдений x_1, \dots, x_n , мы можем принять одно из двух решений: сделать еще хотя бы одно наблюдение или сделать остановку. Допустим, что принято первое из этих решений. Тогда мы будем иметь $n+1$ наблюдений $x_1, \dots, x_n, \xi_{n+1}$. Следуя после этих наблюдений оптимальному правилу остановки, мы понесем убыток, математическое ожидание которого равно $q(x_1, \dots, x_n, \xi_{n+1})$.

Риск после n наблюдений при условии, что будет сделано еще хотя бы одно наблюдение и после этого будет применяться оптимальное правило остановки, равен

$$q_1(x_1, \dots, x_n) = M [q(x_1, \dots, x_n, \xi_{n+1}) | \xi_1 = x_1, \dots, \xi_n = x_n]. \quad (1.136)$$

Предположим, что принято второе решение: остановить процесс наблюдений. Тогда будет понесен убыток $y_n(x_1, \dots, x_n)$.

Чему же равен оптимальный риск? Если $y_n(x_1, \dots, x_n) \leq q_1(x_1, \dots, x_n)$, то выгоднее сделать остановку после n наблюдений, т. е. $q(x_1, \dots, x_n) = y_n(x_1, \dots, x_n)$;

если же $y_n(x_1, \dots, x_n) > q_1(x_1, \dots, x_n)$, выгоднее сделать одно наблюдение, так что $q(x_1, \dots, x_n) = q_1(x_1, \dots, x_n)$. Объединение этих двух случаев приводит к рекуррентной формуле для оптимального риска:

$$q(x_1, \dots, x_n) = \min \{ y_n(x_1, \dots, x_n), M [q(x_1, \dots, x_n, \xi_{n+1}) | \xi_1 = x_1, \dots, \xi_n = x_n] \}. \quad (1.137)$$

Зная $q(x_1, \dots, x_{n+1})$ при всех x_{n+1} , мы можем по этой формуле вычислить $q(x_1, \dots, x_n)$.

Оптимальное правило остановки теперь определяется очень просто: если получены n наблюдений $\xi_1 = x_1, \dots, \xi_n = x_n$, вычисляем значения $q_1 = q_1(x_1, \dots, x_n)$ и $y_n = y_n(x_1, \dots, x_n)$; если $y_n < q_1$, обязательно делаем остановку; если $y_n > q_1$, обязательно делаем еще хотя бы одно наблюдение; если, наконец, $y_n = q_1$, оба решения — остановка и продолжение наблюдений — равнозначны в смысле риска, и можно принять любое из них.

Однако, любая рекуррентная формула требует некоторых начальных условий: нужно каким-то образом задать $q(x_1, \dots, x_N)$, и тогда можно будет вычислить $q(x_1, \dots, x_n)$ при всех $n < N$. Эффективный метод такого задания состоит в следующем.

Заранее ограничим число наблюдений некоторым N и будем выбирать оптимальное правило остановки $v^{(N)}$ при этом ограничении. Отметив все введенные выше функции при данном ограничении индексом (N) вверху, получим:

$$\left. \begin{aligned} q^{(N)}(x_1, \dots, x_n) &= \min \{ y_n(x_1, \dots, x_n), \\ &M [q^{(N)}(x_1, \dots, x_n, \xi_{n+1}) | \xi_1 = x_1, \dots, \xi_n = x_n] \} \quad (n < N); \\ q^{(N)}(x_1, \dots, x_N) &= y_N(x_1, \dots, x_N). \end{aligned} \right\} (1.138)$$

Таким образом, процесс вычисления $q^{(N)}(x_1, \dots, x_n)$ вполне определен. Правило остановки $v^{(N)}$ определяется, как и выше, функциями $y_n(x_1, \dots, x_n)$ и $q_1^{(N)}(x_1, \dots, x_n)$.

Заметим теперь, что для оптимального правила остановки v

$$I(v) = q,$$

где q — частный случай $q(x_1, \dots, x_n)$ при $n=0$.

Далее $I(v^{(N)}) = q^{(N)}$.

Очевидно, $q \leq q^{(N)}$. Если при достаточно большом N

$$q^{(N)} - q \leq \epsilon,$$

где $\epsilon > 0$ — допустимое отклонение от оптимального риска, то вместо v можно использовать $v^{(N)}$.

Таким образом, следует лишь найти соответствующее N , если, конечно, оно существует. Современная теория подобных вопросов изложена в книге [11]. Здесь же мы ограничимся следующим рассуждением.

Пусть

$$y_n = y_n(x_1, \dots, x_n) = C_n + \gamma(x_1, \dots, x_n),$$

где $0 \leq \gamma(x_1, \dots, x_n) \leq K$.

Тогда, если $\beta = P(v > N)$, то

$$q^{(N)} - q \leq \beta K.$$

В то же время при любом $n \geq 0$

$$\beta \leq \frac{M[y_n]}{C_{N+1}}.$$

Следовательно,

$$q^{(N)} - q \leq \frac{1}{C_{N+1}} K \inf_{n \geq 0} M[y_n] \leq \frac{1}{C_{N+1}} K^2.$$

Таким образом, если $C_n \rightarrow \infty$ при $n \rightarrow \infty$, то по любому $\epsilon > 0$ найдется соответствующее значение N .

Несмотря на принципиальную простоту построения оптимального правила остановки, фактическое вычисление оптимального риска представляет собой трудную задачу. Поэтому представляет интерес рассмотрение случаев, когда построение оптимального правила остановки может быть осуществлено непосредственно. Один из таких случаев рассматривается ниже. При выполнении определенных условий монотонности, как следует из леммы, которая будет приведена, возможно непосредственное определение оптимального управления процессом.

Пусть имеется возможность наблюдать последовательность случайных величин $\xi_1, \xi_2, \dots, \xi_n, \dots$ (Природа этих величин не играет никакой роли; в частности, это могут быть векторы многомерного пространства). Совместный закон распределения этих случайных величин предполагается известным. Пусть, далее задана функция убытка $y_n = y_n(x_1, \dots, x_n)$ ($n \geq 0$), имеющая согласно предыдущему смысл убытка при остановке процесса на

n -м шаге после получения наблюдений со значениями $\xi_1 = x_1, \dots, \xi_n = x_n$.

Лемма¹. Пусть существует правило остановки $v^*(x_1, x_2, \dots)$, обладающее следующими свойствами:

1. $M[v^*]$ существует и конечно.

2. Для любых $n \geq 0$ и x_1, \dots, x_n выполняется неравенство

$$M[y_{n+1}(x_1, \dots, x_n, \xi_{n+1})] \leq y_n(x_1, \dots, x_n)$$

при $n \leq v^*(x_1, \dots, x_n, \dots)$ и неравенство

$$M[y_{n+1}(x_1, \dots, x_n, \xi_{n+1})] \geq y_n(x_1, \dots, x_n)$$

при $n \geq v^*(x_1, \dots, x_n, \dots)$.²

3. Для некоторой константы K

$$M[|y_{n+1}(x_1, \dots, x_n, \xi) - y_n(x_1, \dots, x_n)|] \leq K$$

при любых n, x_1, \dots, x_n .

При указанных свойствах v^* является оптимальным правилом остановки среди всех правил остановки v , для которых $M[v]$ конечно.

Заметим, что в соответствии с нашим основным определением, оптимальность правила остановки понимается в смысле минимизации функционала $I(v)$. Если же задана функция пользы и указанный функционал надлежит максимизировать, лемма перефразируется симметричным образом. Условие 2 леммы в этом случае заменится следующим.

2'. Для любых $n \geq 0$ и x_1, \dots, x_n выполняется неравенство

$$M[y_{n+1}(x_1, \dots, x_n, \xi_{n+1})] \geq y_n(x_1, \dots, x_n)$$

при $n \leq v^*(x_1, \dots, x_n)$ и неравенство

$$M[y_{n+1}(x_1, \dots, x_n, \xi_{n+1})] \leq y_n(x_1, \dots, x_n)$$

при $n \geq v^*(x_1, \dots, x_n, \dots)$.

Для доказательства леммы установим сначала следующее утверждение, назвав его «свойством усекаемости».

¹ См. работы [16, 55, 56].

² В силу определения правила остановки выполнение неравенств $n \leq v(x_1, \dots, x_n, \dots)$ и $n \geq v(x_1, \dots, x_n, \dots)$ не зависит от переменных x_{n+1}, x_{n+2}, \dots .

Обозначение $M[y_{n+1}(x_1, \dots, x_n, \xi_{n+1})]$ означает условное математическое ожидание $y_{n+1}(\xi_1, \dots, \xi_n, \xi_{n+1})$ при $(\xi_1, \dots, \xi_n) = (x_1, \dots, x_n)$.

Пусть ν — произвольное правило останковки. Фиксируем некоторое $N \geq 0$ и зададим правило останковки $\nu^{(N)}$ следующим образом. Если для последовательности $x_1, x_2, \dots, x_n, \dots$ правило $\nu(x_1, \dots, x_n, \dots) \leq N$, то $\nu^{(N)}(x_1, \dots, \dots, x_n, \dots) = \nu(x_1, \dots, x_n)$; если же $\nu(x_1, \dots, x_n, \dots) > N$, то $\nu^{(N)}(x_1, \dots, x_n) = N$. В этом случае $\nu^{(N)}$ называется *усеченным правилом останковки*.

Свойство возможности усечения. Пусть выполняется условие 3 леммы, а ν — произвольное заданное правило останковки, для которого $M[\nu] < \infty$. Тогда $|I(\nu)| < \infty$ и для любого $\varepsilon > 0$ можно указать такое N , что

$$I(\nu^N) \leq I(\nu) \pm \varepsilon.$$

Таким образом, если правило останковки усекать достаточно далеко, математическое ожидание основного функционала y , претерпит от этого сколь угодно малое изменение.

Доказательство. Введем символ $E(A)$ для обозначения случайной величины, принимающей значение «1» при происшествии события A и значение «0» в противном случае. Тогда, например, $E(\nu=0)$ есть случайная величина, которая равна единице при $\nu=0$ и равна нулю при $\nu \neq 0$. В таком случае

$$I(\nu) = \sum_{n=0}^{\infty} M[y_n] E(\nu=n),$$

поэтому

$$|I(\nu)| \leq \sum_{n=0}^{\infty} M[|y_n|] E(\nu=n).$$

Если $\nu=0$, утверждение леммы выполняется очевидным образом; поэтому предположим, что $\nu > 0$. В этом случае без ограничения общности можно считать, что $y_0=0$, и предыдущая формула переписется так:

$$|I(\nu)| \leq \sum_{k=1}^{\infty} M[|y_k - y_{k-1}|] \sum_{n=k}^{\infty} \sum_{k=1}^n M[|y_k - y_{k-1}|] E(\nu=n).$$

После изменения порядка суммирования

$$|I(\nu)| \leq \sum_{k=1}^{\infty} M[|y_k - y_{k-1}|] \sum_{n=k}^{\infty} E(\nu=n).$$

Однако, как легко видеть

$$\sum_{n=k}^{\infty} E(v=n) = E(v \geq k).$$

Тогда

$$|I(v)| \leq \sum_{k=1}^{\infty} M[|y_k - y_{k-1}|] E(v \geq k).$$

Заметим теперь, что событие $\{v \geq k\}$ зависит только от $\xi_1, \xi_2, \dots, \xi_{k-1}$. Это событие означает, что процесс наблюдения не остановлен после $(k-1)$ -го шага.

Можно записать

$$M[|y_k - y_{k-1}|] E(v \geq k) = P(v \geq k) M[|y_k - y_{k-1}| | v \geq k].$$

Так как событие $\{v \geq k\}$ зависит только от ξ_1, \dots, ξ_{k-1} , а при любых фиксированных значениях этих переменных условное математическое ожидание значения $|y_k - y_{k-1}|$ не больше K , то

$$M[|y_k - y_{k-1}| | v \geq k] \leq K.$$

Просуммировав по всем k , найдем

$$|I(v)| \leq K \sum_{k=1}^{\infty} P(v \geq k) = KM[v] < \infty.$$

Таким образом, первое утверждение «свойства возможности усечения» доказано.

Рассмотрим теперь правило остановки v^N . Имеем

$$I(v^N) = \sum_{n=0}^{N-1} M[y_n] E(v=n) + M[y_N] E(v \geq N).$$

Вспомнив, что $I(v) = \sum_{n=0}^{\infty} M[y_n] E(v=n)$, приходим к следующей формуле:

$$\begin{aligned} I(v) - I(v^N) &= \sum_{n=N}^{\infty} M[y_n - y_N] E(v=n) = \\ &= \sum_{n=N}^{\infty} \sum_{k=N+1}^{\infty} M[(y_k - y_{k-1})] E(v=n) = \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=N+1}^{\infty} M[(y_k - y_{k-1})] \sum_{n=k}^{\infty} E(v=n) = \\
&= \sum_{k=N+1}^{\infty} M[(y_k - y_{k-1})] E(v \geq k). \quad (1.139)
\end{aligned}$$

На основании полученного выше результата последнее выражение представляет собой остаток сходящегося ряда. Отсюда $I(v^N) \rightarrow I(v)$.

Сформулируем теперь некоторую лемму (назовем ее «частной леммой»).

Частная лемма. Пусть выполнены условия леммы. Тогда для любого N справедливо неравенство

$$I(v^{*N}) \leq I(v^N)$$

при любом правиле остановки v^i .

Докажем, что из частной леммы следует лемма. Пусть v — любое правило остановки с конечным значением $M\{v\}$. Тогда при любом N $I(v^{*N}) \leq I(v^N)$. В то же время $I(v^{*N}) \rightarrow I(v)$, $I(v^N) \rightarrow I(v)$. Отсюда $I(v^*) \leq I(v)$.

Докажем частную лемму. Для простоты изложения ограничимся случаем, когда ξ_n могут принимать лишь конечное число различных значений. Общее доказательство можно осуществить по тому же принципу, однако лишь с помощью абстрактной записи некоторых соотношений, что усложнило бы понимание существа вопроса.

Допустим, что v — некоторое правило остановки, для которого $v \leq N$, причем v^N не совпадает с v^{*N} .

Существуют две, не исключающие одна другую возможности.

1. Может иметь место такой набор $x_1^{(0)}, \dots, x_n^{(0)}$ ($n < N$), что

$$v^{*N}(x_1^{(0)}, \dots, x_n^{(0)}, \dots) > n, \text{ а } v(x_1^{(0)}, \dots, x_n^{(0)}, \dots) = n.$$

2. Может иметь место такой набор $x_1^{(0)}, \dots, x_n^{(0)}$ ($n \leq N$), что

$$v^{*N}(x_1^{(0)}, \dots, x_n^{(0)}, \dots) < n, \text{ а } v(x_1^{(0)}, \dots, x_n^{(0)}, \dots) = n.$$

Рассмотрим первую возможность. Определим новое решающее правило μ следующим образом:

$$\mu(x_1^{(0)}, \dots, x_n^{(0)}, \dots) = n + 1;$$

если $(x_1, \dots, x_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})$, то

$$\mu(x_1, \dots, x_n, \dots) = v^{(*N)}(x_1, \dots, x_n, \dots).$$

По формуле полного математического ожидания

$$I(v) = P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})) M[y_n | (x_1^{(0)}, \dots, x_n^{(0)})] + \\ + P((\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})) M[y_n | (\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})].$$

С другой стороны

$$I(\mu) = P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})) M[y_{n+1} | (x_1^{(0)}, \dots, x_n^{(0)})] + \\ + P((\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})) M[y_{n+1} | (\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})].$$

Отсюда

$$I(\mu) - I(v) = M[y_{n+1} | (x_1^{(0)}, \dots, x_n^{(0)}) - y_n | (x_1^{(0)}, \dots, x_n^{(0)})] \times \\ \times P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})).$$

Ввиду того, что $n < v(x_1^{(0)}, \dots, x_n^{(0)}, \dots)$, по условию 2 леммы последнее выражение меньше или равно нулю. Это означает, что если согласно правилу остановки v процесс наблюдения останавливается раньше, чем это предписывается правилом остановки, определенным леммой, то можно определить новое правило остановки, не худшее предыдущего и такое, что для рассматриваемого значения вектора наблюдений $(x_1^{(0)}, \dots, x_n^{(0)})$ решение об остановке принимается на один шаг позднее. Очевидно, эту процедуру можно осуществить для произвольного вектора $(x_1^{(0)}, \dots, x_n^{(0)})$, для которого

$$v(x_1^{(0)}, \dots, x_n^{(0)}, \dots) = n \text{ и } v^{*N}(x_1^{(0)}, \dots, x_n^{(0)}) > 0.$$

Приходим к выводу, что, не увеличив значения $I(v)$, можно полностью исключить существование векторов указанного вида.

Рассмотрим теперь вторую возможность. Пусть существуют векторы $(x_1^{(0)}, \dots, x_n^{(0)})$, для которых

$$v^{*N}(x_1^{(0)}, \dots, x_n^{(0)}, \dots) \leq n; \quad v(x_1^{(0)}, \dots, x_n^{(0)}, \dots) > n + 1.$$

Выберем из них вектор, для которого n максимально (такой вектор найдется, поскольку по определению $v^{*N}(x_1, x_2, \dots) \leq N$ для любых x_1, x_2, \dots). Это означает, что $v(x_1^{(0)}, \dots, x_n^{(0)}, z) = n + 1$ для любого z .

По формуле полного математического ожидания

$$I(v) = P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})) M[y_{n+1} | (x_1^{(0)}, \dots, x_n^{(0)}, \xi_{n+1})] + \\ + P((\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})) M[y_{n+1} | (\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})].$$

Определим новое правило остановки μ следующим образом:

$$\mu(x_1^{(0)}, \dots, x_n^{(0)}, \dots) = n;$$

если $(x_1, \dots, x_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})$, то

$$\mu(x_1, \dots, x_n, \dots) = v^{*N}(x_1, \dots, x_n, \dots). \quad (1.140)$$

Имеем

$$I(\mu) = P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})) y_n(x_1^{(0)}, \dots, x_n^{(0)}) + \\ + P((\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})) M[y_{n+1}(\xi_1, \dots, \xi_n) \neq (x_1^{(0)}, \dots, x_n^{(0)})]$$

Вычтя из этого выражения предыдущее, получим

$$I(\mu) - I(v) = P((\xi_1, \dots, \xi_n) = (x_1^{(0)}, \dots, x_n^{(0)})) \{y_n(x_1^{(0)}, \dots, x_n^{(0)}) - \\ - M[y_{n+1}(x_1^{(0)}, \dots, x_n^{(0)}, \xi_{n+1})]\}.$$

По условию 2 леммы $I(\mu) - I(v) \leq 0$.

Итак, правило остановки v можно заменить правилом остановки μ , для которого решение об остановке принимается ранее. Последовательно применив эту процедуру нужное число раз, мы исключим и вторую возможность. В результате придем к правилу остановки $v^{(*N)}$, удовлетворяющему условию 2 леммы.

НАДЕЖНОСТЬ АДАПТИВНЫХ СИСТЕМ УПРАВЛЕНИЯ ПОЛЕТОМ

2.1. СТАБИЛИЗАЦИЯ И УПРАВЛЕНИЕ КАК ПРОЦЕССЫ ОБСЛУЖИВАНИЯ «НЕТЕРПЕЛИВЫХ КЛИЕНТОВ»

Эффективность ЛА в значительной мере определяется надежностью, точностью и качеством управления. Усложнение задач, решаемых многорежимными ЛА, влечет за собой непрерывный рост требований к этим характеристикам.

Удовлетворение сформулированным требованиям шло по нескольким линиям и, в частности, по пути создания новых типов высокоточных датчиков информации, разработки адаптивных автопилотов и включения бортовых цифровых вычислительных машин (БЦВМ) в контуры управления ЛА [1, 25, 39, 40, 67, 68].

Существует представление о том, что дополнительное оборудование, необходимое для придания системе свойства приспособляемости к изменениям внутренней и внешней среды, приводит к снижению надежности системы относительно внезапных отказов. При этом игнорируется тот факт, что самонастраивающиеся системы (СНС) с иерархически упорядоченной структурой компенсируют довольно широкий класс внутренних неисправностей основного контура, включая отдельные изменения структуры алгоритма управления. Экспериментальные исследования показали, что такие автоматы в принципе способны компенсировать отдельные отказы, которые приводили к вырождению дифференциального уравнения основного контура управления [34].

Для обоснованного выбора одного из конкурирующих вариантов адаптивного автомата, удовлетворяющего заданным требованиям по надежности, уже на ранних этапах проектирования необходимо иметь адекватные математические модели надежности системы (с учетом

эффекта самоприспособления) и располагать методикой расчета характеристик безотказности.

Какие же черты поведения и структуры иерархически упорядоченных адаптивных систем следует отнести к существенным для построения математической модели их надежности?

Поскольку поведение адаптивной системы находится в непосредственной связи с меняющимся окружением, в конструируемой модели должны быть представлены:

— статистические свойства внутренней и внешней среды (интенсивности отказов компонентов системы с учетом нагрузок и спектральные плотности внешних возмущений);

— способность к частичной компенсации параметрических и структурных нарушений;

— числовые характеристики случайного времени адаптации;

— многоконтурность.

Из этого следует, что математическая модель надежности адаптивной системы должна отражать не только результат адаптации (достижение цели в меняющемся окружении), но и некоторые важные черты самого процесса адаптации, в частности, запаздывание ее по отношению к изменениям внешней обстановки, наблюдаемое во всех СНС с обратной связью по критерию качества.

Модель должна быть чувствительной к вариациям существенных параметров. Исследование модели должно привести к эффективному вычислению интересующих нас характеристик безотказности адаптивной системы:

— вероятности безотказной работы за время полета (или за какой-либо другой период);

— математического ожидания времени до попадания системы в поглощающее состояние (до полного отказа);

— математического ожидания времени до первого некомпенсированного параметрического отказа.

Успешное построение математической модели во многом определяется удачным выбором метатеории модели. В качестве метатеории мы воспользуемся теорией массового обслуживания (ТМО). По мере надобности будут использоваться также методы теории регулирования и теории случайных процессов.

Процессы адаптации — переходные процессы во внешних контурах СНС — мы будем рассматривать с позиций ТМО как процессы обслуживания «нетерпеливых клиен-

тов»¹, к которым относятся такие явления ожидания, когда на время пребывания клиента в очереди или в обслуживающем приборе накладываются ограничения. В интересующих нас случаях время ожидания должно быть равно нулю. Иначе говоря, поступившая в случайный момент времени заявка на адаптацию должна немедленно обслуживаться. В терминах теории регулирования это означает, что случайный момент возникновения ошибки или рассогласования является случайным моментом зарождения переходного процесса в системе.

Разумеется, процессы отработки рассогласования между желаемым и действительным поведением в обычных (несамонастраивающихся) системах также могут быть изучены с позиций ТМО, и мы сначала обратимся к простому примеру [26].

В системе «самолет-автопилот» (рис. 2.1) автопилот обеспечивает заранее заданный режим полета (например, поддерживает неизменным заданный курс), а при подаче внешних команд заставляет самолет выполнять различные эволюции. Регулируемой координатой здесь является курс ψ . Желаемое значение курса ψ_0 задается первоначальной уставкой наружного кольца курсового гироскопа Г, сохраняющего неизменной свою ориентацию в пространстве в условиях рыскания самолета. В случайный момент времени сигнал рассогласования $\Delta\psi$ снимается с потенциометра П, движок которого связан с наружным кольцом гироскопа, а основание потенциометра — с корпусом самолета.

Действительно, при указанном вращении основания потенциометра между его движком и средней точкой возникает рассогласование $\Delta\psi$, которое в соответствии с выбранным законом управления преобразуется автопилотом и через рулевую поверхность и внешнюю среду создает управляющий момент. Последний заставляет самолет поворачиваться так, чтобы рассогласование $\Delta\psi$ все время удерживалось близким к нулю.

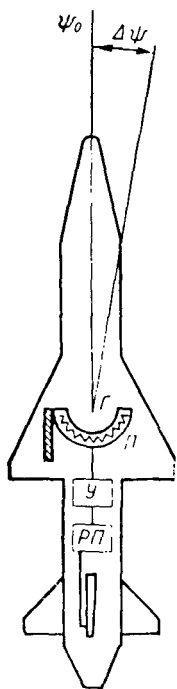


Рис. 2.1. Простейшая одноконтурная система «самолет-автопилот»

Если предусмотреть устройство (например, в виде червячной пары), позволяющее вращать основание потенциометра относительно корпуса самолета, то появляется возможность подавать команды на изменение курса. При работе автопилота в определенных режимах желаемое значение курса может изменяться во времени в соответствии с программой полета.

¹ Термин «нетерпеливые клиенты» был введен в 1956 г. Д. У. Баррером [51, 52].

В описанной одноконтурной системе имеются все основные части СМО: входящий поток клиентов, обслуживающий прибор, и выходящий поток обслуженных клиентов. Источником входящего потока является сравнивающее устройство, реализованное в датчике углового положения. Момент прихода «нетерпеливого клиента» — это случайный момент выхода отклонения курса $\Delta\varphi$ за уровни, определяемые требованиями к точности стабилизации или управления. Обслуживающим прибором является усилительно-преобразовательное устройство У автопилота, рулевой привод РП и сам ЛА. С выхода аппарата снимается поток обслуженных заявок. Обратная связь такой СМО замыкается через аэродинамику объекта. Потеря «нетерпеливого клиента», т. е. возникновение нового рассогласования в период отработки команды или ошибки, может грозить потерей устойчивости или управляемости самолета.

В соответствии со сложившейся в теории массового обслуживания традицией мы начнем изучение СМО с анализа статистической структуры входящего потока, затем перейдем к изучению процесса обслуживания, построению и решению системы уравнений и, наконец, к проверке адекватности модели существенным чертам изучаемого явления.

Говоря об адаптивных системах, мы будем иметь в виду многоконтурные самонастраивающиеся системы.

Все реальные адаптивные системы обладают ограниченной приспособляемостью, и раньше всего следует уточнить ее пределы, т. е. конкретизировать существенную черту — способность к частичной компенсации параметрических и структурных нарушений.

2.2. ПРИСПОСОБЛЯЕМОСТЬ АДАПТИВНЫХ СИСТЕМ К СОБСТВЕННЫМ ЧАСТИЧНЫМ НЕИСПРАВНОСТЯМ

Все наблюдаемые изменения параметров в непрерывных системах стабилизации и управления можно подразделить на [61]:

- α -вариации, не приводящие к изменению начальных условий и порядка уравнения;
- β -вариации, или изменения начальных условий;
- λ -вариации, приводящие к изменению порядка дифференциального уравнения движения системы.

К настоящему времени накоплен обширный экспериментальный материал, свидетельствующий о способности контуров самонастройки (КС) в избыточных адаптивных системах компенсировать α -, β - и в ограниченной степени λ -вариации параметров основного контура.

При некоторых ограничениях на относительную сложность контура самонастройки возможно достижение свое-

го рода «равновесия» между дополнительной ненадежностью, привносимой КС, и расширением компенсаторных возможностей адаптивной системы как целостной структуры. Иначе говоря, суммарная интенсивность внезапных отказов элементов основного контура, компенсируемых за счет деятельности исправного КС, может быть сделана близкой к суммарной интенсивности внезапных отказов собственно КС.

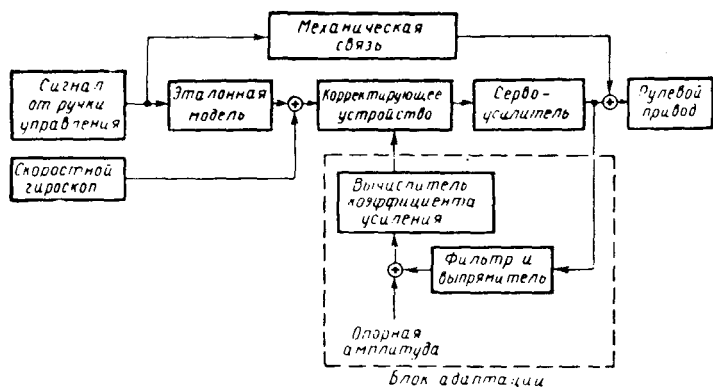


Рис. 2.2. Обобщенная структурная схема адаптивного автопилота МН-96

При этом надо иметь в виду, в частности то, что:

— компенсируются поломки преимущественно тех элементов, которые участвуют в формировании настраиваемого параметра;

— при адаптации используется информация об отказах, содержащаяся в вызванных ими возмущениях.

Ниже приводятся результаты электронного моделирования и летных испытаний самонастраивающегося автопилота МН-96 (рис. 2.2), установленного на ракетоплане Х-15 [28, 58, 60, 69, 70]. Экспериментальный аэрокосмический аппарат Х-15 предназначен для суборбитальных полетов и отработки парящего входа в атмосферу Земли спускаемых космических аппаратов. Хотя с начала летных испытаний Х-15 прошло более 10 лет, результаты этих полетов продолжают привлекать внимание специалистов [69].

Испытывалось несколько аппаратов Х-15, причем некоторые полеты совершались с обычными автопилотами

с постоянным коэффициентом усиления. Это обстоятельство позволило провести детальное сравнение поведения обычных автопилотов с адаптивной системой МН-96 в широком диапазоне режимов (в атмосферном и внеатмосферном полетах, а также при входе в атмосферу). Коэффициент усиления аппарата Х-15 в ходе полетов изменялся примерно в 1000 раз.

С целью снижения перегрузок и уменьшения нагрева аппарата возвращение с больших высот происходило при больших углах атаки. Коэффициенты статической устойчивости при больших углах атаки существенно отличаются от коэффициентов для небольших углов атаки при одних и тех же скорости и высоте полета.

Обычный автопилот с программным изменением угла атаки по скоростному напору не может обеспечить желаемую степень устойчивости объекта на всех режимах полета. Способность СНС приспосабливаться к изменению аэродинамических характеристик оказалась важной для обеспечения хорошей управляемости в процессе возвращения аппарата на Землю.

Результаты моделирования полета аппарата Х-15 на больших высотах с использованием ручного реактивного управления и управления с помощью автопилота МН-96 показали, что при адаптивной автоматической стабилизации расход топлива в два раза меньше, чем при ручном управлении. Летные испытания Х-15 свидетельствуют о том, что преимущества адаптивного автопилота МН-96 проявляются при внеатмосферных полетах и на участке входа в атмосферу.

Рассмотрим вопросы надежности адаптивной системы. В результате электронного моделирования и испытаний реального образца автопилота МН-96 выяснилось, что 90—95% отказов его электронного оборудования составляли отказы типа «обрыв» («О») и «короткое замыкание» («КЗ»), к которым автопилот не приспосабливался.

Отказы компонентов, приводившие к уменьшению коэффициента усиления, компенсировались контуром адаптации. В действующем варианте автопилота использовались интегральные схемы. Все приборы МН-96 были задублированы. Для того, чтобы раздельное параллельное дублирование было эффективным, каждый прибор снабжался надежным предохранителем, который сводил отказы типа «КЗ» к отказам типа «О», т. е. отключал вышедший из строя прибор. Таким образом обеспечи-

валась защищенность автопилота МН-96 от любого одиночного и некоторых двойных внезапных отказов.

Предусматривался также предстартовый контроль: на авиаматке Б-52, с которой производится запуск ракетоплана Х-15, устанавливался автоматический анализатор, который проверял более 70 параметров и функций автопилота МН-96. Цикл проверки занимал около трех минут [60].

Эквивалентная средняя наработка на отказ электронного оборудования автопилота МН-96 равна около 100 000 ч для двухчасового полета. За первые два года эксплуатации только один элемент автопилота МН-96 вышел из строя. Этот отказ не ухудшил качества функционирования, а вызвал лишь незначительное отклонение угла рыскания, которое пилот ощутил как слабое нарушение курсовой балансировки. За первые 850 ч эксплуатации автопилота (суммарная наработка) зафиксировано семь неисправностей, причем пять из них вызваны ошибками оператора.

По более поздним данным [69] на ракетоплане Х-15 было совершено всего 65 полетов, во время которых зафиксировано два уже упомянутых выше внезапных отказа автопилота МН-96. Последствия этих отказов, возникших в одном из каналов, были локализованы и не распространились на другие каналы.

За время летной эксплуатации в автопилоте МН-96 наблюдалось несколько сбоев, последствия которых иногда оказывали влияние на другие каналы автопилота. Благодаря исправно действовавшему контуру самонастройки все сбои были скомпенсированы за время переходного процесса, вызванного возмущением.

При анализе летных испытаний [60, 69, 70] отмечается хорошее качество технического обслуживания.

Катастрофа, происшедшая с третьим аппаратом Х-15 (Х-15-3), требует тщательного рассмотрения¹. Для разбора обстоятельств этого полета необходимо вновь остановиться на некоторых подробностях работы контура самонастройки системы МН-96.

Коэффициент усиления системы должен быть обратно пропорционален эффективности рулей. Контур самонастройки управляется амплитудой колебаний, происходящих на собственной частоте сервопривода. Узкополосный

¹ Анализ последнего полета ракетоплана Х-15-3 основан на материалах, опубликованных в работах [58, 69, 70].

фильтр КС должен обеспечить подавление колебаний на всех частотах, не совпадающих с собственной частотой сервопривода. Выходной сигнал фильтра подается на вход выпрямителя. С выхода выпрямителя снимается напряжение постоянного тока, пропорциональное амплитуде колебаний на его входе. Этот сигнал сравнивается с фиксированной опорной амплитудой. Отличное от нуля рассогласование приводит в действие исполнительный механизм контура самонастройки.

Теоретически КС должен реагировать только на колебания с желаемой собственной частотой сервопривода. В действительности же почти все случайные колебания содержат в своем спектре частот составляющие, которые проходят через узкополосный фильтр. Поэтому такие факторы, как турбулентность, упругие колебания ЛА, электрические помехи и управляющие сигналы от штурвала оказывают некоторое влияние на КС. Это влияние в большинстве случаев сводится к тому, чтобы «звести в заблуждение» КС и представить дело таким образом, как будто бы коэффициент усиления в прямой цепи слишком высокий и его необходимо уменьшить до величины, ниже критической.

Контур самонастройки может быть «введен в заблуждение» также сигналами постоянного тока или низкочастотными сигналами, которые оказываются достаточно мощными, чтобы насытить электрические допуски и замаскировать сигналы с частотами, лежащими в пределах полосы пропускания фильтра. В этих случаях коэффициент усиления будет возрастать до значений, превышающих критические, и сервопривод потеряет устойчивость. Амплитуда неустойчивых колебаний будет возрастать до тех пор, пока скорость выходной координаты сервопривода не достигнет своего предельного значения. В этот момент возникнут колебания предельного цикла и они будут поддерживаться. Такого рода ситуация и возникла в полете.

Поведение ЛА и характеристики его управляемости в значительной степени зависят от того, каким способом автопилот с усилением сигнала от ручки управления «подсоединен» к самолету. В большинстве пилотируемых аппаратов адаптивные системы с усилением управляющего сигнала (здесь сервоприводу отводится ограниченная роль) соединяются последовательно с основной механической или гидравлической системой управления. При

таким характере соединения ограничения на команду не имеют особого значения, поскольку пилот через основную систему управления может усилить управляющее воздействие. Такие команды приведут к немедленному насыщению электронной системы усиления входного сигнала. Если к тому же эти управляющие сигналы достаточно мощные и неоправданы обстоятельствами, то последствия потенциально могут быть опасными для нормальной деятельности контура самонастройки. Очевидное, но не всегда практически реализуемое решение заключается в разработке и установке на ЛА полностью электрической системы управления, исключаяющей прямую связь пилота с поверхностями управления.

Известная опасность кроется также в использовании одних и тех же аэродинамических рулей для адаптивной стабилизации аппарата по каналам тангажа и крена. В этих аппаратах необходимость демпфирования или угловой стабилизации по одной из осей будет мешать эффективной работе другого канала. Если бы произошло насыщение в одном из каналов, то это привело бы к полной потере управляемости по другому каналу.

Отмеченные недостатки присущи не только адаптивным системам с большим коэффициентом усиления. Однако они усугубляются, когда методология построения таких систем используется для решения задач устойчивости и управляемости ЛА с адаптивными автопилотами. Если случится, что все упомянутые явления возникнут одновременно, то последствия могут стать катастрофическими.

Телеметрическая информация о последнем полете Х-15-3 включает в себя примеры работы адаптивного автопилота МН-96, коэффициент усиления которой находился в некритической области. Кроме того, телеметрия показывает, как контур самонастройки может быть «введен в заблуждение».

На рис. 2.3 представлены осциллограммы некоторых характеристик высотного полета аппарата Х-15-3 с запланированным потолком 76 000 м. Здесь представлены как функции времени: высота полета H , углы тангажа θ и крена γ , перемещение цилиндра сервопривода правого горизонтального стабилизатора L и уровень адаптивного коэффициента усиления по каналу тангажа K_θ . Как правило, он дает некоторое представление о значениях коэффициентов усиления в остальных каналах.

Система работала нормально после старта примерно в течение 60 с. На 61-й секунде в аппарате возникла электрическая помеха, которая затронула работу всех систем. Эта помеха, вероятно, наводилась экспериментом, проводившимся в одной из подвесок крыла. Хотя

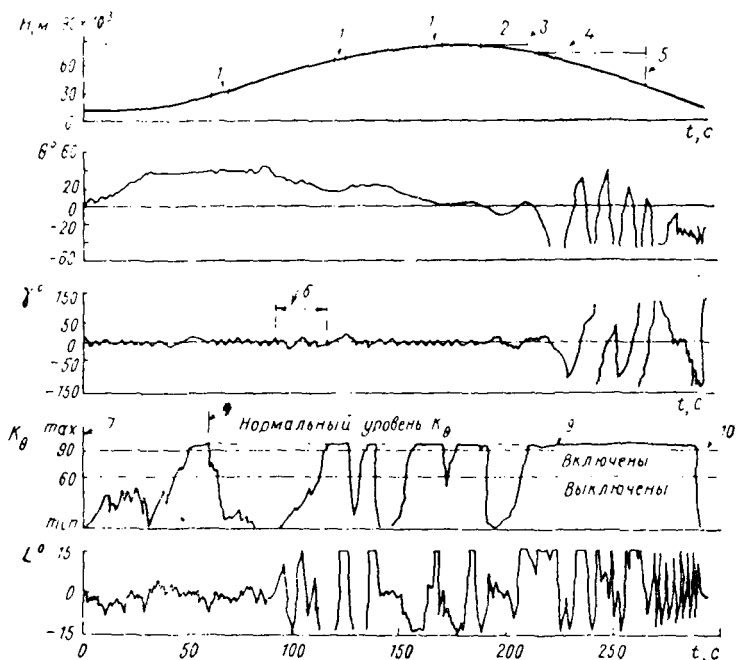


Рис. 2. 3. К анализу полета аппарата X-15-3:

1—включение нормальных реактивных органов; 2—включение ручного реактивного управления; 3—вхождение в штопор; 4—включение нормальных реактивных органов; 5—выход из штопора; 6—маневры по каналу крена; 7—запуск; 8—начало воздействия электрической помехи; 9—прекращение действия помехи; 10—автоматические реактивные органы

адаптивная система управления полетом была изолирована, помеха оказала существенное влияние на ее работу, в основном, на деятельность контура адаптации. На всем протяжении полета, пока существовала помеха, коэффициенты усиления часто находились ниже критического уровня и создавались препятствия их перестройке для достижения более высоких значений. Электрическая помеха в какой-то момент достигла столь высокого уровня, что привела к разъединению рулевых сервопри-

водов в каналах тангажа и крена. Потребовалось вмешательство пилота, который восстановил их взаимодействие за 5 с.

Во время атмосферного полета влияние такой помехи на передаточные числа не повлекло бы столь серьезных последствий, поскольку даже при минимальных коэффициентах усиления характеристики управляемости остаются вполне приемлемыми. Однако, для рассматриваемого высотного полета воздействие электрической помехи оказалось существенным, так как оно нарушило работу реактивных органов управления.

В аппарате X-15 реактивные органы применяются в сочетании с аэродинамическими рулями, причем и те и другие управляются из кабины летчика одним регулятором для облегчения пилотирования на больших высотах. «Нормальные» реактивные органы включаются и отключаются автоматически в зависимости от среднего значения коэффициента усиления системы для экономии топлива. Уровни переключения для запуска и остановки реактивного управления были определены в результате летных испытаний (90 и 60% соответственно, как показано на осциллограмме — см. рис. 2.3).

Вскоре после достижения максимальной высоты пилот, по-видимому, осознал наличие неполадок в системе управления, так как начал использовать левый регулятор, который приводит в действие двойное ручное реактивное управление. Стабилизируя полет с помощью реактивных органов, он, очевидно, неправильно истолковал показания приборов и начал изменять курс, хотя это и не было оправдано обстоятельствами.

Реактивные органы управления препятствовали бы этому изменению, если бы они были исправны. Пилот продолжал изменять угол рыскания аппарата, пока он не отклонился на 90° от курса на высоте около 72 960 м, идя со снижением. В этот момент аппарат вошел в штопор.

С целью выхода из штопора пилот использовал реактивные и аэродинамические рули. Нормальные реактивные органы были также автоматически включены и работали почти непрерывно, препятствуя возрастанию скорости штопора.

Примерно через 10 с после вхождения в штопор перестал работать демпфер канала рыскания в результате команды от системы обнаружения неисправностей. Эта

система должна обнаруживать отказавший гироскоп, но в данном случае она была «введена в заблуждение» и автоматически отключила исправный курсовой скоростной гироскоп.

Проектом было предусмотрено, что при превышении скоростью рыскания величины 20 град/с выход скоростного гироскопа по этому каналу отключается. Это и было сделано. Однако пилот не получил никаких сведений об отключении гироскопа.

В результате совместных действий органов управления и собственной аэродинамической устойчивости самолета штопор прекратился на высоте около 46 380 м. Во время штопора все передаточные числа достигли своих предельных значений, поскольку все «нормальные» сигналы, которые могли бы уменьшить коэффициенты усиления, были замаскированы мощными управляющими сигналами от пилота и большими угловыми скоростями.

Когда самолет вышел из штопора, начала развиваться неустойчивость сервопривода в канале тангажа; причиной могли быть либо сигналы от штурвала, либо управляющие воздействия автопилота. Они мешали КС уменьшать коэффициент усиления в канале тангажа, что он обязан делать при возрастании скоростного напора.

Поверхности горизонтального стабилизатора поворачивались симметрично на своей предельной скорости 26 град/с, совершая пилообразные колебания с полной амплитудой, равной 20°. При таком действии рулевых поверхностей возникли и поддерживались колебания аппарата по каналу тангажа. Как только рулевой привод потерял устойчивость, пилот больше не мог эффективно управлять или поддерживать запас устойчивости ни в канале тангажа, ни в канале крена (поверхности горизонтального стабилизатора используются для управления и по тангажу и по крену). В результате недемпфированных движений аппарата Х-15-3 возникли перегрузки, превысившие его прочностные характеристики.

Из проведенного после этого полета моделирования как будто следовало, что уменьшение коэффициента усиления, осуществленное адаптивной системой или пилотом, или отключение автопилота могли прекратить эти колебания. По какой-то неизвестной причине, может быть из-за отсутствия информации о неполадках в автопилоте, или действия больших перегрузок, пилот ничего не предпринял, чтобы с помощью ручных органов управления

ения, и аппарат рассыпался, м [69].

— неоспоримое средство деятельности адаптивной системы к сбоям. Однако, его поставках стадиях проектирования, макеты или опытные образцы, необходимость определения структуры системы ощущается на проектировании при анализе математических вариантов адаптивной теории автоматического регулирования различные методы, позволяющие частности, метод частотных характеристик годографа и др.

кой квазистационарной системы выходя условий параметрической инвариантная функция основного контура системы [15]

$$\frac{K}{s) + bG(s) + Q(s)}, \quad (2.1)$$

зи движения объекта; главный параметр основного контура равен параметра a . В теории полюсов замкнутой системы, функцией (2.1), при одновременном

ы имеет вид

$$(s) + Q(s) = 0. \quad (2.2)$$

ажение плоскости независимых параметров $\delta + j\omega$.

$$\left. \begin{aligned} &+ a \operatorname{Re} H(s); \\ &+ a \operatorname{Im} H(s); \end{aligned} \right\} \quad (2.3)$$

у

$$-Q(s). \quad (2.4)$$

ормное отображение вспомогательности траекторий полюсов $s = \delta + j\omega$. Взаимосвязи изменений одного параметра другого является совпадением направлений на плоскости s . На плоскости

по отдельное рассмотрение процессов в траектории, рассматриваются в п. 2.5.

(a, b) эти векторы являются ортогональными. Поскольку отображение $z \rightarrow s$ является конформным, изменение угла между векторами $a = \text{const}$ и $b = \text{const}$ возможно только за счет преобразования (2.3).

Векторы, ортогональные на плоскости (a, b), совпадают по направлению на плоскости z (а следовательно, и на плоскости s), если

$$\text{Im } G(s) \text{Re } H(s) - \text{Im } H(s) \text{Re } G(s) = 0. \quad (2.5)$$

Это условие является необходимым и достаточным для полной компенсации.

Выражения для корневых годографов замкнутой системы имеют вид:

$$\left. \begin{aligned} 1. \quad a \text{ var: } & \text{Im} [Q(s) + bG(s)] \text{Re } H(s) - \\ & - \text{Im } H(s) \text{Re} [Q(s) + bG(s)] = 0; \\ 2) \quad b \text{ var: } & \text{Im} [Q(s) + aH(s)] \text{Re } G(s) - \\ & - \text{Im } G(s) \text{Re} [Q(s) + aH(s)] = 0. \end{aligned} \right\} \quad (2.6)$$

Совместное решение уравнений (2.5) и (2.6) дает координаты точек, в которых выполняется условие полной компенсации.

К сожалению, вывод необходимого и достаточного условий полной компенсации (параметрической инвариантности) в замкнутой аналитической форме для систем с передаточной функцией произвольного вида связан с большими трудностями и пока не получен. С теорией параметрической компенсации более глубоко можно ознакомиться в работе [38].

2.3. СТРУКТУРА ВХОДЯЩЕГО ПОТОКА

Рассмотрим двухконтурную непрерывную самонастраивающуюся систему с m параметрами настройки $(\alpha_1, \dots, \alpha_m)$, состояние которой характеризуется m -мерным стационарным гауссовским процессом $V(t)$ (рис. 2.4). На рисунке $g(t)$ — управляющий вход; $u(t)$ — возмущения, действующие на объект; $y(t)$ — выходная координата объекта. Выход траектории этого процесса за пределы некоторой заранее очерченной области может быть отождествлен с потерей качества регулирования, а в случае монотонных траекторий — с потерей устойчивости.

Пересечения допустимой области траекториями процесса $V(t)$ определяют моменты времени, когда корректирующее воздействие контура адаптации на параметрические входы основного контура становится необходимым. Если траектории — монотонные неубывающие функции, то такие воздействия необходимы для сохранения устойчивости системы.

Из-за различных требований к расчету регулирования на различных участках траектории полета граница допустимой области в общем случае является функцией

времени. Таким образом, нас будут интересовать вопросы, связанные с распределением во времени пересечений выборочных функций однородных по времени гауссовских процессов с заданными постоянными или криволинейными уровнями. Упомянутые моменты времени образуют поток случайных событий — параметрических

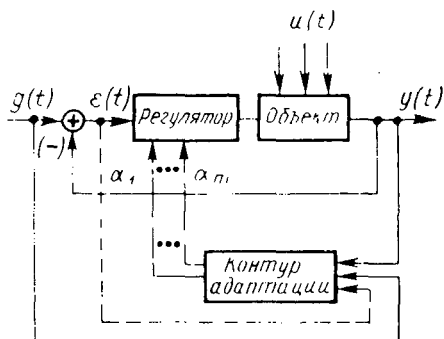


Рис. 2.4. Обобщенная структурная схема двухконтурной бесперебойной СИС

отказов — с неизвестным заранее законом распределения

Непрерывное опознавание состояния основного контура позволяет контуру адаптации контролировать и корректировать параметрические отказы в латентный (скрытый) период, предшествующий их внешнему проявлению — потере качества регулирования. (Заметим, что в замкнутом адаптивном автомате корректирующее воздействие контура самонастройки начинается не в момент пересечения траекторий процесса $V(t)$ с этой границей, а несколько раньше, при достижении рассогласованием $\varepsilon_{к.с}(t)$ порога чувствительности канала самонастройки).

Необходимость в самонастройке возникает в связи с изменениями как внешней, так и внутренней среды. Поэтому поток компенсируемых параметрических отказов (заявок на адаптацию) образуется в результате наложения потоков из нескольких источников: внутренних и внешних.

В качестве меры надежности автомата относительно параметрических отказов выберем математическое ожидание числа таких событий в единицу времени. Возника-

ет естественный вопрос: если существует некоторый класс отказов, который автоматически диагностируется и корректируется, то зачем вводить меру надежности автомата относительно этих явлений? Ответ на этот вопрос заключается в том, что существуют быстротекущие управляемые процессы, которые при некоторых неисправностях могут попасть в поглощающее состояние раньше, чем адаптивный регулятор сможет справиться с новой, возникшей в результате отказа ситуацией.

Математическое ожидание числа пересечений m -мерного стационарного гауссовского процесса с многомерной поверхностью S за время T можно определить по формуле Беляева [7]. Если процесс одномерный, т. е. $m=1$ и уровень $u=u(t)$ — дифференцируемая функция, то применяется формула Лидбеттера [59]

$$M[N_{V(t)}; u(t); 0, T] = \frac{1}{2\sqrt{q(0)}} \int_0^T f\left(\frac{u(t)}{\sqrt{q(0)}}\right) \times \\ \times \left\{ 2\sqrt{-q''(0)} f\left(\frac{u'(t)}{\sqrt{-q''(0)}}\right) + u'(t) \left[2\Phi\left(\frac{u'(t)}{\sqrt{-q''(0)}}\right) - 1 \right] \right\} dt, \quad (2.7)$$

где $q(\tau)$ — дважды дифференцируемая нормированная корреляционная функция процесса $V(t)$;

$f(y)$ — плотность нормального распределения;

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt — функция Лапласа.$$

При постоянном уровне значение $M[N_{V(t)}]$ определяется по формуле Булинской [13]; эта формула является частным случаем выражения (2.7) при $m=1$, $u=\text{const}$:

$$M[N_{V(t)}; u; 0, T] = \frac{T}{2\pi} \left(-\frac{q''(0)}{q(0)} \right)^{\frac{1}{2}} e^{-\frac{u^2}{2q(0)}}. \quad (2.8)$$

В дальнейшем мы будем полагать, что параметрические отказы образуют не однородный по времени пуассоновский поток. Если выборочные функции процесса — гладкие кривые с ограниченными производными, то это допущение хорошо согласуется с действительностью. Это относится, в частности, к таким случайным процессам, как коэффициент усиления ЛА или эффективность его

аэродинамических рулей. Как известно, их выборочные функции — почти всюду гладкие кривые. Скачкообразное изменение их траекторий возможно лишь при весьма неточной начальной установке параметров или сбросе отработавших ступеней ЛА.

Мгновенную интенсивность потока заявок на адаптацию будем определять по известной формуле

$$\bar{v}(t) = \frac{dM\{N_{V(t)}; 0, t\}}{dt} \quad (2.9)$$

2.4. ПОЛУМАРКОВСКИЙ ПРОЦЕСС В АДАПТИВНЫХ СИСТЕМАХ

В непрерывных двухконтурных адаптивных системах, рассмотрением которых мы ограничимся, нас будут интересовать следующие случайные потоки:

— не однородный по времени пуассоновский поток параметрических отказов с мгновенной интенсивностью $\bar{v}(t)$;

— однородный пуассоновский поток внезапных отказов в основном контуре с интенсивностью η ;

— однородный пуассоновский поток внезапных отказов в контуре самонастройки с интенсивностью λ .

Изменения свойств окружающей среды, разумеется, не происходят скачками, но для того, чтобы иметь возможность ввести количественную меру *времени адаптации* рассмотрим случай, когда требуется дискретное изменение уровней настраиваемых параметров. Такое положение в какой-то мере соответствует включению автопилота ЛА на произвольном режиме полета при неизвестных заранее значениях параметров.

Процесс изменения внутренней среды автомата непрерывный, но на его фоне возможны отдельные ступенчатые выбросы. Пользуясь терминологией А. Н. Колмогорова, такой процесс можно назвать процессом с дискретным вмешательством случая.

Таким образом в адаптивных системах протекает смешанный случайный процесс.

Случайность разрывных по величине или производным возмущающих воздействий и начальных условий порождает случайность переходных процессов в адаптивных системах. Будем полагать, что:

— время адаптации является случайной величиной с произвольным законом распределения $F(u) = P\{t_{ад} < u\}$;

— самонастройка начинается в момент достижения выборочной функцией случайного процесса $V(t)$ границы области S ;

— процесс адаптации квазистационарный, т. е. «очередь» параметрических отказов (заявок на обслуживание) не образуется.

Случайный процесс развития состояний адаптивной системы будем изучать с помощью фазового пространства, состоящего из множества изолированных точек и полупрямой ($0 \leq u < \infty$). При сделанных допущениях эволюция состояний системы образует полумарковский процесс (см. п. 1.3).

Для нас будут представлять интерес различные состояния, образующие полную группу несовместных событий:

— состояние $\omega_0(t)$, характеризующееся тем, что в рассматриваемый момент t в системе отсутствуют параметрические и внезапные отказы;

— состояние $\omega_1(t, u)$, характеризующееся тем, что в момент t контур самонастройки занят компенсацией параметрического отказа в течение времени u , и внезапные отказы в основном контуре (ОК) и контуре самонастройки (КС) не возникли;

— состояние $\omega_2(t)$, характеризующееся тем, что в момент t возник внезапный отказ в КС, но отсутствуют внезапные и параметрические отказы в ОК; в этом состоянии возможна ложная, т. е. не оправданная состоянием ОК адаптация;

— состояние $\omega_3(t)$, характеризующееся тем, что в момент t произошел параметрический отказ в ОК, неисправен КС, но внезапные отказы в ОК не возникли;

— состояние $\omega_4(t)$, характеризующееся тем, что в момент t система полностью неработоспособна.

Состояниям $\omega_0(t)$, $\omega_2(t)$, $\omega_3(t)$, $\omega_4(t)$ в фазовом пространстве соответствуют изолированные точки; состоянию $\omega_1(t, u)$ — полупрямая ($0 \leq u < \infty$). (Граф переходов адаптивной системы стабилизации приведен в п. 2.6 на рис. 2.9).

Перейдем к выводу системы интегро-дифференциальных уравнений, описывающей случайный процесс развития состояний системы при произвольном законе распределения времени адаптации. Определим вероятности пребывания системы в каждом из состояний в момент $t + \Delta t$ при условии, что известна вероятность пребывания в

этом состоянии в момент t . В результате мы должны получить систему пяти уравнений с пятью неизвестными

$$P_0(t), P_1(t), P_2(t), P_3(t), P_4(t):$$

$$1) P_0(t + \Delta t) = P_0(t) [1 - \pi(t, t + \Delta t)] (1 - \lambda \Delta t) (1 - \eta \Delta t) + \\ + \int_0^t p_1(t, u) \frac{F(u + \Delta t) - F(u)}{1 - F(u)} du + o(\Delta t),$$

где $\pi(t, t + \Delta t)$ — вероятность того, что в промежутке $(t, t + \Delta t)$ возникает хотя бы один параметрический отказ;

$1 - \lambda \Delta t + o(\Delta t)$ — вероятность отсутствия внезапных отказов в КС;

$1 - \eta \Delta t + o(\Delta t)$ — вероятность отсутствия внезапных отказов в ОК;

$P_0(t)$ — вероятность пребывания в состоянии $\omega_0(t)$ в момент t ;

$p_1(t, u)$ — плотность вероятности (по u) того, что в момент t заявка обслуживается и к этому моменту самонастройка длится время u_0 , $u < u_0 < u + \Delta u$;

$\frac{F(u + \Delta t) - F(u)}{1 - F(u)}$ — вероятность того, что адаптация закончилась за промежуток $(t, t + \Delta t)$;

$$2) p_1(t + \Delta t, u + \Delta t) = p_1(t, u) \frac{1 - F(u + \Delta t)}{1 - F(u)} (1 - \lambda \Delta t) \times \\ \times (1 - \eta \Delta t) + o(\Delta t),$$

где $\frac{1 - F(u + \Delta t)}{1 - F(u)}$ — вероятность того, что адаптация не завершилась за время $(t, t + \Delta t)$;

$$3) P_2(t + \Delta t) = \lambda \Delta t P_0(t) + [1 - \pi(t, t + \Delta t)] (1 - \eta \Delta t) P_2(t) + o(\Delta t);$$

$$4) P_3(t + \Delta t) = \lambda \Delta t \int_0^t p_1(t, u) du + \pi(t, t + \Delta t) P_2(t) + \\ + (1 - \eta \Delta t) P_3(t) + o(\Delta t);$$

$$5) P_4(t + \Delta t) = \eta \Delta t [P_0(t) + \int_0^t p_1(t, u) du] + P_2(t) + \\ + P_3(t) + o(\Delta t).$$

Граничное условие

$$p_1(t, \Delta t) \Delta t = \pi(t, t + \Delta t) (1 - \eta \Delta t - \lambda \Delta t) P_0(t) + o(\Delta t).$$

Начальные условия:

$$P_0(0) = 1; P_2(0) = P_3(0) = P_4(0) = 0.$$

Осуществляя предельный переход при $\Delta t \rightarrow 0$, получим систему интегро-дифференциальных уравнений с переменным коэффициентом $\bar{v} = \bar{v}(t)$:

$$\left. \begin{aligned} \dot{P}_0(t) &= -(\bar{v}(t) + \lambda + \eta) P_0(t) + \int_0^t p_1(t, u) \mu(u) du; \\ \frac{\partial p_1(t, u)}{\partial t} + \frac{\partial p_1(t, u)}{\partial u} &= -[\mu(u) + \lambda + \eta] p_1(t, u); \\ \dot{P}_2(t) &= \lambda P_0(t) - (\bar{v}(t) + \eta) P_2(t); \\ \dot{P}_3(t) &= \lambda \int_0^t p_1(t, u) du + \bar{v}(t) P_2(t) - \eta P_3(t); \\ \dot{P}_4(t) &= \eta \left[\int_0^t p_1(t, u) du + \sum_{\substack{i=0 \\ (i \neq 1)}}^3 P_i(t) \right]. \end{aligned} \right\} (2.10)$$

Граничное условие

$$p_1(t, 0) = \bar{v}(t) P_0(t). \quad (2.11)$$

Начальные условия имеют прежний вид.
Условие нормировки:

$$\int_0^t p_1(t, u) du + \sum_{\substack{i=0 \\ (i \neq 1)}}^4 P_i(t) = 1. \quad (2.12)$$

Для решения системы (2.10) введем замену

$$p_1^*(t, u) = \frac{p_1(t, u)}{1 - F(u)}, \quad (2.13)$$

после чего система (2.10) примет вид:

$$\left. \begin{aligned} \dot{P}_0(t) &= -[\bar{v}(t) + \lambda + \eta] P_0(t) + \int_0^t p_1^*(t, u) g(u) du; \\ \frac{\partial p_1^*(t, u)}{\partial t} + \frac{\partial p_1^*(t, u)}{\partial u} &= -(\lambda + \eta) p_1^*(t, u); \\ \dot{P}_2(t) &= \lambda P_0(t) - (\bar{v}(t) + \eta) P_2(t); \end{aligned} \right\} (2.14)$$

$$\left. \begin{aligned} \dot{P}_3(t) &= \lambda \int_0^t p_1^* [1 - F(u)] du + \bar{v}(t) P_2(t) - \eta P_3(t) \\ \dot{P}_4(t) &= \eta \left\{ \int_0^t p_1^* [1 - F(u)] du + \sum_{\substack{i=0 \\ (i \neq 1)}}^3 P_i(t) \right\}. \end{aligned} \right\} (2.14)$$

Система (2.14) может быть решена численными методами, например, методом сеток. При $\bar{v}(t) = \bar{v} = \text{const}$ решение получится в замкнутой аналитической форме, если применить преобразование Лапласа и известные теоремы операционного исчисления (о свертке и др.).

Опуская промежуточные выкладки, запишем решение системы (2.14) в изображениях:

$$\left. \begin{aligned} \bar{P}_0(s) &= \{s + \lambda + \eta + \bar{v}[1 - \bar{g}(s + \lambda + \eta)]\}^{-1}; \\ \bar{P}_1(s) &= \bar{P}_0(s) \frac{\bar{v}[1 - \bar{g}(s + \lambda + \eta)]}{s + \lambda + \eta}; \\ \bar{P}_2(s) &= \bar{P}_0(s) \frac{\lambda}{s + \bar{v} + \eta}; \\ \bar{P}_3(s) &= \bar{P}_0(s) \frac{\bar{v}\lambda}{s + \eta} \left[\frac{1 - \bar{g}(s + \lambda + \eta)}{s + \lambda + \eta} + \frac{1}{s + \bar{v} + \eta} \right]; \\ \bar{P}_4(s) &= \eta \bar{P}_0(s) \left\{ 1 + \frac{\bar{v}[1 - \bar{g}(s + \lambda + \eta)]}{s + \lambda + \eta} + \frac{\lambda}{s + \bar{v} + \eta} + \right. \\ &\quad \left. + \frac{\bar{v}\lambda}{s + \eta} \left[\frac{1 - \bar{g}(s + \lambda + \eta)}{s + \lambda + \eta} + \frac{1}{s + \bar{v} + \eta} \right] \right\}. \end{aligned} \right\} (2.15)$$

Для системы, способной функционировать (по крайней мере на отдельных участках траектории полета) с неисправным, но отключенным (во избежание ложной адаптации) контуром самонастройки, множество благоприятных состояний

$$\Omega = \{\omega_i\}, \quad i = 0, 1, 2, 3, \quad (2.16)$$

Следовательно, вероятность безотказной работы системы в изображениях

$$\bar{P}_c(s) = \sum_{i=0}^3 \bar{P}_i(s). \quad (2.17)$$

Математическое ожидание времени до первого внезапного отказа (полной поломки) системы можно определить из выражения (2.9), не прибегая к обратному преобразованию Лапласа, а именно:

$$T_0 = \bar{P}_c(s) \Big|_{s=0} = [\lambda + \eta + \bar{v}(1 - \bar{g}(\lambda + \eta))]^{-1} \times \\ \times \left\{ 1 - \frac{\bar{v}}{\lambda + \eta} [1 - \bar{g}(\lambda + \eta)] + \frac{\lambda}{\bar{v} + \eta} + \frac{\bar{v}\lambda}{\eta} \left[\frac{1 - \bar{g}(\lambda + \eta)}{\lambda + \eta} + \frac{1}{\bar{v} + \eta} \right] \right\}. \quad (2.18)$$

Если безотказная работа контура адаптации необходима на всей траектории полета, то множество благоприятных состояний образуется из $\omega_0(t)$, $\omega_1(t)$, $\omega_2(t)$. В этом случае вероятность безотказной работы адаптивной системы в изображениях

$$\bar{P}_c(s) = \sum_{i=0}^2 \bar{P}_i(s). \quad (2.19)$$

Математическое ожидание времени до первого некомпенсированного параметрического отказа двухконтурной адаптивной системы равно

$$T_n = \sum_{i=0}^2 \bar{P}_i(s) \Big|_{s=0} = [\lambda + \eta + \bar{v}(1 - \bar{g}(\lambda + \eta))]^{-1} \times \\ \times \left\{ 1 + \frac{\bar{v}[1 - \bar{g}(\lambda + \eta)]}{\lambda + \eta} + \frac{\lambda}{\bar{v} + \eta} \right\}. \quad (2.20)$$

При известном законе распределения времени самонастройки, применяя обратное преобразование Лапласа, можно получить характеристики безотказности адаптивной системы в оригиналах.

В важном для приложений частном случае экспоненциального закона распределения времени адаптации $g(t) = \mu e^{-\mu t}$:

$$\bar{g}(s + \lambda + \eta) = \int_0^{\infty} e^{-(s+\lambda+\eta)t} \mu e^{-\mu t} dt = \frac{\mu}{s + \mu + \lambda + \eta}. \quad (2.21)$$

В этом случае:

$$\bar{P}_0(s) = \frac{s + \mu + \lambda + \eta}{(s + \bar{v} + \lambda + \eta)(s + \mu + \lambda + \eta) - \bar{v}\mu}; \quad (2.22)$$

$$\begin{aligned} \bar{P}_1(s) &= \frac{\nu}{(s + \bar{\nu} + \lambda + \eta)(s + \mu + \lambda + \eta) - \bar{\nu}\mu}; \\ \bar{P}_2(s) &= \frac{\lambda(s + \mu + \lambda + \eta)}{(s + \bar{\nu} + \eta)[(s + \bar{\nu} + \lambda + \eta)(s + \mu + \lambda + \eta) + \bar{\nu}\mu]}; \\ \bar{P}_3(s) &= \frac{\bar{\nu}\lambda[2(s + \eta) + \bar{\nu} + \lambda + \mu]}{(s + \eta)(s + \bar{\nu} + \eta)[(s + \bar{\nu} + \lambda + \eta) \times \\ &\quad \times (s + \mu + \lambda + \eta) - \bar{\nu}\mu]}; \\ \bar{P}_4(s) &= \frac{\bar{\nu}[\eta(s + \eta)(s + \bar{\nu} + \eta) + \lambda\eta(s + \bar{\nu} + \eta)] - \\ &\quad - (s + \mu + \lambda + \eta)[- \bar{\nu}\lambda\eta - \eta(s + \eta)(s + \bar{\nu} + \eta) - \lambda\eta(s + \eta)]}{s(s + \eta)(s + \bar{\nu} + \eta)[(s + \bar{\nu} + \lambda + \eta) \times \\ &\quad \times (s + \mu + \lambda + \eta) - \bar{\nu}\mu]}. \end{aligned}$$

Применив обратное преобразование Лапласа, чим:

$$\begin{aligned} P_0(t) &= \frac{\mu}{\bar{\nu} + \mu} e^{-(\lambda + \eta)t} + \frac{\bar{\nu}}{\bar{\nu} + \mu} e^{-(\bar{\nu} + \mu + \lambda + \eta)t}; \\ P_1(t) &= \frac{\bar{\nu}}{\bar{\nu} + \mu} e^{-(\lambda + \eta)t} - \frac{\bar{\nu}}{\bar{\nu} + \mu} e^{-(\bar{\nu} + \mu + \lambda + \eta)t}; \\ P_2(t) &= \frac{\lambda(\lambda + \mu - \nu)}{\lambda(\lambda + \mu - \nu) - \bar{\nu}\mu} e^{-(\bar{\nu} + \eta)t} + \\ &\quad + \frac{\lambda\mu}{(\bar{\nu} - \lambda)(\bar{\nu} + \mu)} e^{-(\lambda + \eta)t} - \frac{\bar{\nu}\lambda}{(\mu + \lambda)(\mu + \bar{\nu})} e^{-(\bar{\nu} + \mu + \lambda + \eta)t}; \\ P_3(t) &= e^{-\eta t} - \frac{\lambda(\mu + \lambda - \nu)}{\lambda(\mu + \lambda - \bar{\nu}) - \bar{\nu}\mu} e^{-(\bar{\nu} + \eta)t} - \\ &\quad - \frac{\bar{\nu}(\bar{\nu} + \mu - \lambda)}{(\bar{\nu} - \lambda)(\bar{\nu} + \mu)} e^{-(\lambda + \eta)t} + \frac{\bar{\nu}\lambda}{(\mu + \lambda)(\bar{\nu} + \mu)} e^{-(\bar{\nu} + \mu + \lambda + \eta)t}; \\ P_4(t) &= 1 - e^{-\eta t}. \end{aligned}$$

Если время адаптации распределено по закон номерной плотности, т. е.

$$g(t) = \begin{cases} \frac{1}{b-a}, & a \leq t \leq b; \\ 0, & t > b, t < a, \end{cases}$$

то

$$\begin{aligned} \bar{g}(s | \lambda + \eta) &= \int_0^{\infty} e^{-(s+\lambda+\eta)t} [\gamma(t-a) - \gamma(t-b)] \frac{1}{b-a} dt = \\ &= e^{-a(s+\lambda+\eta)} [1 - e^{-(b-a)(s+\lambda+\eta)}] \frac{1}{(s+\lambda+\eta)(b-a)}, \quad (2.25) \end{aligned}$$

где $\gamma(t)$ — единичная функция.

Если при этом условия полета предусматривают необходимость деятельности КС на всей траектории, то математическое ожидание времени до первого попадания системы в поглощающее состояние $\omega_3(t)$ равно

$$\begin{aligned} T_n &= \sum_{i=0}^2 \bar{P}_i(s) \Big|_{s=0} = \\ &= \left\{ \lambda + \eta + \bar{v} \left[1 - \frac{e^{-a(\lambda+\eta)} [1 - e^{-(b-a)(\lambda+\eta)}]}{(\lambda+\eta)(b-a)} \right] \right\}^{-1} \times \\ &\times \left\{ 1 + \frac{\bar{v}}{\lambda+\eta} \left[1 - \frac{e^{-a(\lambda+\eta)} [1 - e^{-(b-a)(\lambda+\eta)}]}{(\lambda+\eta)(b-a)} \right] + \frac{\lambda}{\bar{v} + \eta} \right\}. \quad (2.26) \end{aligned}$$

При экспоненциальном законе распределения времени адаптации

$$\begin{aligned} T_n &= \frac{1}{\lambda + \eta} + \frac{\lambda(\lambda + \mu - \bar{v})}{[\lambda(\lambda + \mu - \bar{v}) - \bar{v}\mu](\bar{v} + \eta)} + \\ &+ \frac{\lambda\mu}{(\bar{v} - \lambda)(\bar{v} + \mu)(\lambda + \eta)} - \frac{\bar{v}\lambda}{(\mu + \lambda)(\mu + \bar{v})(\bar{v} + \mu + \lambda + \eta)}. \quad (2.27) \end{aligned}$$

Из последней формулы нетрудно усмотреть, как влияют на T_n параметры λ , η , \bar{v} , μ .

На рис. 2.5 приводятся графики зависимости T_n от скорости процесса адаптации для различных значений мгновенной интенсивности \bar{v} потока параметрических отказов. При бесконечно большой скорости адаптации ($\mu \rightarrow \infty$) эффект влияния параметрических отказов сводится к минимуму, кривая функции $T_n = f(\mu)$ асимптотически

$$\text{стремится к } \max T_n = \frac{1}{\lambda + \eta} \left[1 + \frac{\lambda}{\bar{v} + \eta} \right].$$

При синтезе самонастраивающихся систем на стремление повышать быстродействие КС накладываются ограничения, связанные с необходимостью обеспечения его устойчивости. Приведенные выше формулы и графики (см. рис. 2.5) дают дополнительный материал для суждения о необходимом

конечном быстродействии контура самонастройки в целях обеспечения достаточно высоких субоптимальных значений T_n при фиксированных λ , η , ν .

Реальные самона-

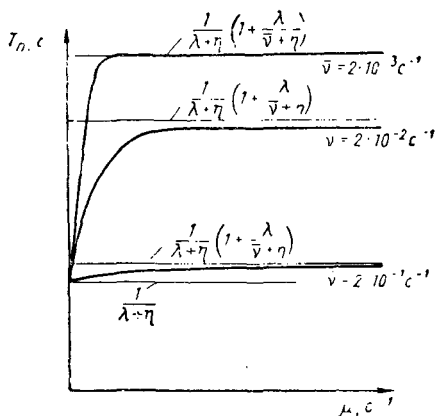


Рис. 2.5. Графики зависимости математического ожидания времени до первого параметрического отказа от скорости процесса адаптации

страивающиеся системы обладают ограниченной приспособляемостью: они не способны компенсировать любые рассогласования или косвенно влиять на вынужденную составляющую переходного процесса от возмущений любой величины. Поэтому значение вероятностей состояний целесообразно использовать в качестве части исходных данных для оценки эффективности системы:

$$E_c(t) = \sum_{i \in \mathcal{Q}} P_i(t) e_i(t), \quad (2.28)$$

где $e_i(t)$ — эффективность адаптивной системы в i -м состоянии в момент t (это может быть, например, условная вероятность выполнения требований технического задания к точности стабилизации).

Таким образом удается отобразить процесс реального функционирования системы с учетом внешних возмущений и изменения работоспособности элементов с течением времени [29—34].

Учет частичной приспособляемости адаптивной системы стабилизации к внезапным отказам элементов ОК (к λ -вариациям вектора внутренних параметров) приводит

к изменению некоторых коэффициентов системы интегродифференциальных уравнений.

Обозначим суммарную интенсивность внезапных отказов, компенсируемых за счет деятельности КС, буквой ξ . Тогда система (2.14) принимает вид:

$$\left. \begin{aligned} \dot{P}_0(t) &= (\bar{v}(t) + \lambda + \eta) P_0(t) + \int_0^t p_1^*(t, u) g(u) du; \\ \frac{\partial p_1^*(t, u)}{\partial t} + \frac{\partial p_1^*(t, u)}{\partial u} &= -[\lambda + (\eta - \xi)] p_1^*(t, u); \\ \dot{P}_2(t) &= \lambda P_0(t) - (\bar{v}(t) + \eta) P_2(t); \\ \dot{P}_3(t) &= \lambda \int_0^t p_1^*[1 - F(u)] du + \bar{v}(t) P_2(t) - \eta P_3(t); \\ \dot{P}_4(t) &= (\eta - \xi) \left[P_0(t) + \int_0^t p_1^*[1 - F(u)] du + \right. \\ &\quad \left. + \eta (P_2(t) + P_3(t)) \right]. \end{aligned} \right\} (2.29)$$

Изменится и граничное условие

$$p_1(t, 0) = [\bar{v}(t) + \xi] P_0(t). \quad (2.30)$$

Алгебраическое суммирование интенсивностей отказов вытекает из того, что пуассоновский поток является самопроизводящим по отношению к операции суперпозиции.

Изложенный здесь подход к построению и исследованию математической модели надежности многоконтурных адаптивных систем может быть использован для аналитического изучения надежности СНС с иной структурной организацией (например, СНС с двухкратной инвариантностью [39] и ультраустойчивых самоорганизующихся систем, основанных на методе избыточных переменных [17]).

Отметим еще возможность сведения немарковского процесса к марковскому путем расширения пространства состояний [21].

Пусть закон распределения случайного времени самонастройки аппроксимируется суперпозицией законов Эрланга различного порядка

$$f(t) = \sum_{l=1}^m c_l \pi(t, k, \mu_l), \quad (2.31)$$

где c_i — постоянные положительные величины, в сумме равные единице;

λ — плотность вероятности закона Эрланга i -го порядка с масштабным параметром μ_i .

Для этого случая простой прием построения марковского процесса с конечным числом состояний заключается в следующем. В качестве состояний процесса достаточно взять пары чисел i, r ($1 \leq i \leq m, 0 \leq r \leq k_i - 1$) и особое состояние 0. В момент достижения состояния 0 разыгрывается случайная ситуация с возможными исходами $1, 2, \dots, m$, с соответствующими вероятностями P_1, P_2, \dots, P_m .

Если имеет место i -й исход, то процесс попадает в состояние $i, 0$. Из состояния i, r ($0 \leq r < k_i - 2$) возможен переход в состояние $i, r+1$, а из состояния $i, k_i - 1$ — возможен переход лишь в состояние 0. Интенсивность перехода в обоих случаях равна μ . (Пример применения этого способа к построению системы уравнений, описывающих эволюцию состояний двухконтурной самонастраивающейся системы с синусоидальным пробным сигналом дан в конце главы).

В результате получается система линейных дифференциальных уравнений Колмогорова с переменными или постоянными коэффициентами. Здесь уместно обратить внимание на одно важное свойство решений уравнений Колмогорова, проявляющееся, когда приведенная плотность потока параметрических отказов $\psi(t) = \frac{\nu(t)}{\mu} > 1$. На временных интервалах, соответствующих $\psi(t) > 1$, темп возникновения параметрических отказов опережает скорость адаптации и наблюдается отличное от нуля расхождение между действительным и желаемым поведением адаптивной системы.

Таким образом, изучение переходных вероятностей дает возможность проследить эффект запаздывания адаптации относительно изменений среды, наблюдаемый при работе адаптивных систем с обратной связью по критерию качества.

2.5. НАДЕЖНОСТЬ КОНТУРА АДАПТАЦИИ

Контур адаптации, решающий ответственные задачи идентификации характеристик ОК и осуществления корректирующих воздействий, должен быть максимально за-

щищен от собственных ошибок: в двухконтурной системе они уже ничем не могут быть скомпенсированы. Это справедливо для любой иерархически упорядоченной системы, поскольку выход из строя управляющего устройства высшего ранга не восполняется деятельностью подчиненных регуляторов. Вопрос о предельно допустимой относительной сложности КС при фиксированной сложности и надежности ОК всегда возникает при синтезе адаптивных систем и представляет далеко не академический интерес.

Введем понятие о правильно организованном контуре самонастройки. Будем называть КС правильно организованным, если он удовлетворяет трем требованиям:

- 1) реагирует на все отказы ОК;
- 2) его собственные неисправности не приводят к тому, что отказавший ОК признается исправным;
- 3) его отказы типа «ложный сигнал» или «короткое замыкание» не приводят к ложной адаптации.

Если КС наделен первыми двумя свойствами, то к устройству идентификации КС применима формула, полученная в 1962 г. С. И. Злочевским, Р. А. Мирным, А. В. Розановым:

$$P_{у.и}(t) = \frac{(P_{о.к}(t) + P_d - 1)}{P_{о.к}(t) P_d}, \quad (2.32)$$

где $P_{у.и}(t)$ — вероятность безотказной работы устройства идентификации за время t ;

$P_{о.к}(t)$ — вероятность безотказной работы ОК за время t ;

P_d — коэффициент доверия (условная вероятность того, что отказ ОК действительно имел место, если об этом сигнализирует устройство идентификации КС).

По формуле (2.32) рассчитана табл. 2.1 с двумя входами.

Из таблицы видно, что при достаточно высоких значениях коэффициента доверия ($P_d \geq 0,9$) вероятность отказа устройства идентификации контура самонастройки должна быть, по крайней мере, на порядок ниже вероятности отказа собственно основного контура (т. е. когда компенсаторные механизмы не действуют).

Таблица 2.1

$P_{о.к}$	$P_{д}$				
	0,9	0,99	0,999	0,9999	1
0,9	0,987	0,998	0,9998	0,99999	1
0,99	0,999	0,9998	0,9999	0	1
0,999	0,9999	0,9999	1	1	1
0,9999	0,99999	1	1	1	1
1	1	1	1	1	1

Если ограничиться рассмотрением объекта контроля и устройства идентификации без структурной избыточности, то из формулы (2.32) и табл. 2.1 можно сделать при фиксированном $P_{д}$ вывод об относительной сложности $C_{у.и}/C_{о.к}$, а именно:

$$\frac{C_{у.и}}{C_{о.к}} = \frac{\Lambda_{у.и}}{\Lambda_{о.к}} \approx 0,1, \quad (2.33)$$

где $\Lambda_{у.и}$, $\Lambda_{о.к}$ — суммарная интенсивность висзапных отказов устройства идентификации и ОК соответственно.

Уместно вспомнить, что в корректирующих кодах, например, в коде Хэмминга с исправлением одиночных ошибок, отношение минимального числа проверочных разрядов k к числу информационных разрядов m ($k \approx \log_2 m$) почти всюду (за исключением коротких слов с $m \leq 3$) является правильной дробью. С увеличением длины двончного слова это отношение убывает, что хорошо прослеживается на графике (рис. 2.6). В основе этого (кажущегося висшним) сходства лежат, быть может, глубокие причины. По-видимому, любая разумная процедура идентификации должна приводить к простым (относительно испытываемого объекта) физическим воплощениям.

Третье требование относится преимущественно к исполнительным устройствам КС и связывается с представлением о необходимости достижения их практической безотказности. Едва ли можно доверять коррекцию устройствам, собственная частота отказов которых сравнима с этой же характеристикой корректируемого объекта. Суммарная интенсивность отказов корректирующих цепей контура самоастройкы должна быть величинной, по

крайней мере, второго порядка малости по сравнению с величиной η . (Справедливости ради заметим, что в нескольких известных нам практических примерах синтеза самонастраивающихся систем это условие оказалось выполненным).

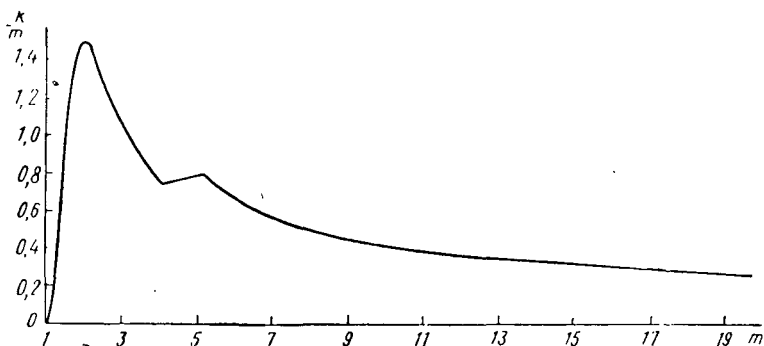


Рис. 2.6. Зависимость отношения минимального числа проверочных разрядов k к числу информационных разрядов m от длины двоичного слова m

Из этих рассуждений следует, что выход приведенной таблицы можно рассматривать как нижнюю грань для вероятности безотказной работы всего КС при заданных $P_{о.к}$ и $P_{д}$. Вопрос о выборе значений $P_{д}$, к сожалению, до сих пор остается нерешенным. Интуитивное представление связывает уровень коэффициента доверия $P_{д}$ с требуемой вероятностью достижения тех целей, для которых создается адаптивная система стабилизации или управления.

В противном случае действие КС не будет вне подозрений, и возникнет надобность в коррекции самого контура адаптации, т. е. создания еще одного уровня управления.

В связи с изложенным, известный интерес представляют высказывания нейрофизиологов Э. А. Асратяна и П. В. Симонова [4], касающиеся некоторых общих принципов компенсации нарушенных функций в высокоорганизованных биологических системах: «...Многоэтажная регуляция не предполагает примитивного дублирования. Каждый новый уровень приносит в регулируемую деятельность новые свойства, новое качество. Подобный принцип обладает, по меньшей мере, двумя следствиями. Подчиненные регуляторы способны функционировать в случае выхода из строя «командующих» механизмов.

В свою очередь, высшие интеграторы освобождаются от элементарных функций для более сложной деятельности.

И опять-таки этот принцип должен носить динамический характер. Высший регулятор может не вмешиваться в работу соподчиненных ему узлов. Если же где-то появляется угроза дисфункции, верхние органы защиты информируются о поломке и мобилизуют механизмы ее ликвидации.

Эти высшие регуляторы должны обладать двумя качествами: 1) высокой степенью чувствительности к сдвигам в различных частях сложной системы, ибо вмешательство, предупреждающее перебои в ее работе, эффективно, если оно своевременно; 2) они сами должны быть максимально защищены от возможных поломок, потому что всякая неисправность регуляторов получает широкий отклик в соподчиненных узлах».

Итак, высокая чувствительность к изменениям вектора внутренних параметров ОК и инвариантность относительно собственных частичных неисправностей — вот те качества, которыми должен обладать правильно организованный контур адаптации в самонастраивающихся системах со структурной избыточностью.

Пример 2.1. Расчет вероятности безотказной работы двухконтурной СИС стабилизации канала крена за время полета.

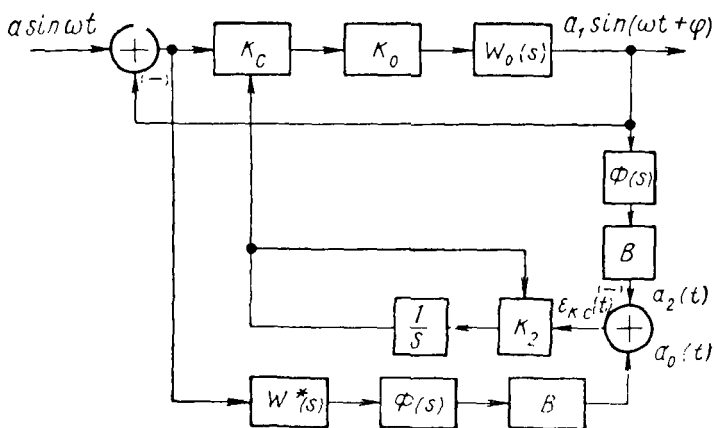


Рис. 2.7. Структурная схема бесконечной СИС с эталонной моделью и контролем амплитудной частотной характеристики

Блок-схема системы стабилизации с эталонной моделью и контролем амплитудной частотной характеристики изображена на рис. 2.7. На вход основного контура от специального генератора пробных сигналов подаются гармонические колебания $a \sin \omega t$. На выходе ОК (рассматривается линейная модель основного контура) возбуждаются вынужденные колебания $a_1 \sin(\omega t + \varphi)$, которые по-

даются на вход контура самонастройки. Последний состоит из резонансного фильтра $\Phi(s)$, выпрямителя B , сравнивающего устройства, усилителя K_2 и интегрирующего звена с передаточной функцией $1/s$. Передаточная функция эталонной модели ОК в разомкнутом состоянии — $W^*(s)$.

Частоту тестовых сигналов выбирают таким образом, чтобы основные спектры управляющих сигналов и возмущений, а также тоны изгибных колебаний ЛА лежали вне полосы пропускания резонансного фильтра. С выхода фильтра $\Phi(s)$ вынужденные колебания подаются на вход выпрямителя B , выход которого $a_2(t)$ сравнивается с желаемым выходным сигналом a_0 . При достижении рассогласованиям $\epsilon_{к.с}(t) = a_0 - a_2(t)$ порога чувствительности канала самонастройки интегрирующее устройство КС осуществляет автоматическое изменение настраиваемого параметра K_c ОК до тех пор, пока произведение $K_c K_0$ не станет равным постоянной величине (с точностью до погрешности и допустимых приборных погрешностей).

Таким образом, стабилизация амплитуды вынужденных колебаний $a_2(t)$ с помощью автоматического изменения K_c означает стабилизацию коэффициента усиления ОК в разомкнутом состоянии.

Заметим, что автоматическая настройка K_c позволяет также компенсировать некоторые вариации постоянной времени объекта.

Динамика систем такого рода подробно исследована, описана в литературе (см., например, [27, 38]) и здесь не рассматривается.

Основные допущения и определения

1. При достаточно медленном изменении параметров объекта и малой скорости процессов настройки основной контур допустимо считать системой с постоянными параметрами, частотные характеристики которой в разомкнутом и замкнутом состояниях

$$W(i\omega) = W(\omega) e^{i\theta W(\omega)} \quad \text{и} \quad \Phi(i\omega) = \Phi(\omega) e^{i\varphi(\omega)}$$

связаны обычным соотношением

$$\Phi(i\omega) = \frac{W(i\omega)}{1 + W(i\omega)}. \quad (2.34)$$

Если к этому добавить предположение о малости ошибок настройки, то процессы настройки в контуре адаптации с точностью до малых величин второго порядка малости могут рассматриваться изолированно от процессов в ОК [27].

2. Вариации коэффициента усиления объекта K_0 происходят вследствие действия на ЛА одномерного нестационарного гауссовского процесса $V(t)$. В процессе $V(t)$ присутствует более или менее ярко выраженная (в зависимости от режима полета) неслучайная составляющая.

Все время полета $t_{пол}$ разбивается на n неперекрывающихся интервалов, в пределах каждого из которых законны допущения о стационарности и нормальности процесса $V(t)$ и известна его спектральная плотность $S_{V_i}(\omega)$, $i=1, 2, \dots, n$.

3. Рассогласование $\epsilon_{к.с}(t)$ в КС в пределах каждого из участков представляет собой одномерный гауссовский стационарный процесс; его спектральная плотность и нормированная корреляционная функция определяются по известным формулам:

$$S_{\epsilon_{к.с}}(\omega) = |W_{\epsilon/V}(i\omega)|^2 S_V(\omega); \quad (2.35)$$

$$\varrho_{\varepsilon_{к.с}}(\tau) = \int_{-\infty}^{\infty} S_{\varepsilon_{к.с}}(\omega) e^{i\omega\tau} d\omega, \quad (2.36)$$

где $W_{\varepsilon/V}(s)$ — передаточная функция КС от возмущения $V(t)$;

$|W_{\varepsilon/V}(i\omega)|$ — модуль частотной характеристики.

4. Самонастройка начинается в момент пересечения выборочной функцией процесса $V(t)$ уровней c_1 и c_2 , соответствующих максимально допустимой ошибке (рис. 2.8). Обслуживание заявки на адаптацию завершается в момент вхождения траектории процесса $\varepsilon_{к.с}(t)$ в допустимые пределы (c_1, c_2).

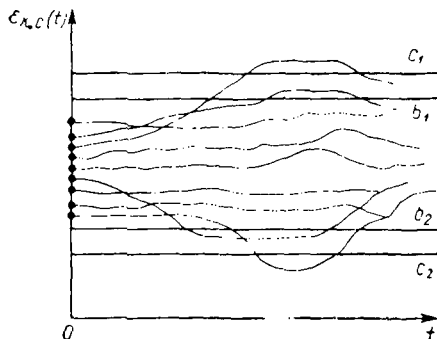


Рис. 2.8. Выбросы случайного процесса $\varepsilon_{к.с}(t)$ за уровни, определяемые требованиями к максимально допустимой ошибке КС

ниями номера участка. Мгновенная интенсивность параметрических отказов определяется для каждого участка по формулам, приведенным в п. 2.3.

7. Отказы элементов ОК и КС образуют однородный по времени пуассоновский поток.

Исходные данные

1. Время работы автомата в полете $\tau_{пол} = 60$ с.
2. Суммарная интенсивность внезапных отказов нерезервированного ОК с учетом электрических, тепловых и механических нагрузок в полете $\eta = 43 \cdot 10^{-7}$ с⁻¹.
3. Суммарная интенсивность внезапных отказов нерезервированного КС с учетом отмеченных нагрузок $\lambda = 15 \cdot 10^{-7}$ с⁻¹.
4. Плотность распределения случайного времени адаптации $\mu = 0,2$ с⁻¹.
5. Кусочно-линейная аппроксимация мгновенной интенсивности параметрических отказов для каждого характерного участка траектории полета сведена в табл. 2.2.
6. Около 40% возможных внезапных отказов элементов ОК компенсируется за счет корректирующего параметрического воздействия. Требуется рассчитать вероятность безотказной работы систем за время полета.

№ участка i	Длина участ- ка, с	$\bar{v}_i(t) = a_i + b_i t, c^{-1}$	№ участка i	Длина участ- ка, с	$\bar{v}_i(t) = a_i + b_i t, c^{-1}$
1	18	$0,0087 - 0,0002t$	5	2,1	$0,1232 - 0,039t$
2	5,4	$0,0055 + 0,0045t$	6	1,2	$0,0412 + 0,2683$
3	3,1	$0,3907 - 0,1265t$	7	6,2	$0,0261$
4	1,6	$0,0621 + 0,0382t$	8	3	$0,0439$
			9	19,4	$0,0194$

Решение

Фазовое пространство адаптивной системы образуется из пяти состояний $\omega_i(t)$, $i=0,1,\dots,4$ (см. п. 2.4). При сделанных предположениях относительно законов распределения случайных потоков развитие состояний системы образует на каждом участке полета не однородный по времени марковский процесс и описывается системой уравнений:

$$\begin{cases}
 \dot{P}_0 = -[a_i \pm b_i t + \lambda + \eta] P_0(t) + \mu P_1(t); \\
 \dot{P}_1 = [a_i \pm b_i t] P_0(t) - [\mu + \lambda + \eta] P_1(t); \\
 \dot{P}_2 = \lambda P_0(t) - [a_i \pm b_i t + \eta] P_2(t); \\
 \dot{P}_3 = \lambda P_1(t) + [a_i \pm b_i t] P_2(t) - \eta P_3(t); \\
 \dot{P}_4 = \eta \sum_{j=1}^3 P_j(t); \quad i = 1, 2, \dots, 9.
 \end{cases} \quad (2.37)$$

Начальные условия: $P_0(0) = 1$, $P_j(0) = 0$, $j=1,\dots,4$.

Соответствующий граф переходов изображен на рис. 2.9. Результаты решения системы (2.37) методом Рунге-Кутты на ЦВМ «Паири» приведены на рис. 2.10.

Из графиков переходных вероятностей $P_0(t)$ и $P_1(t)$ (см. рис. 2.10, а) видно, что наиболее активная работа КС приходится на элементы траектории полета, лежащие в окрестности 32-й секунды. Вероятность гибели автомата $P_4(t)$, несмотря на наличие компенсаторных механизмов, монотонно возрастает с течением времени, хотя и остается допустимо малой в пределах времени полета (см. рис. 2.10, б).

Учет компенсации известной доли внезапных отказов ОК (см. п. 6 основных допущений) исправно действующим КС важен: он приводит к заметному снижению вероятности попадания (кривая $\tilde{P}_4(t)$, рис. 2.10, б) в поглощающий экран $\omega_4(t)$.

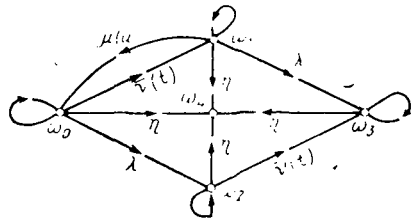


Рис. 2.9. Граф переходов двухконтурной СНС

Вероятности состояний при $\tau_{пол} = 60$ с: $P_0(60) = 0,909238$; $P_1(60) = 0,090231$; $P_2(60) = 0,000034$; $P_3(60) = -0,000055$; $P_4(60) = 0,000470$.

Вероятность безотказной работы СНС за время полета $P_c(60) = 0,999503$.

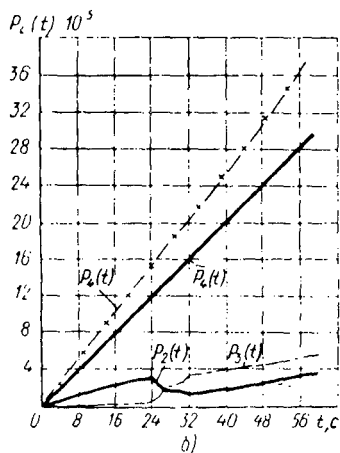
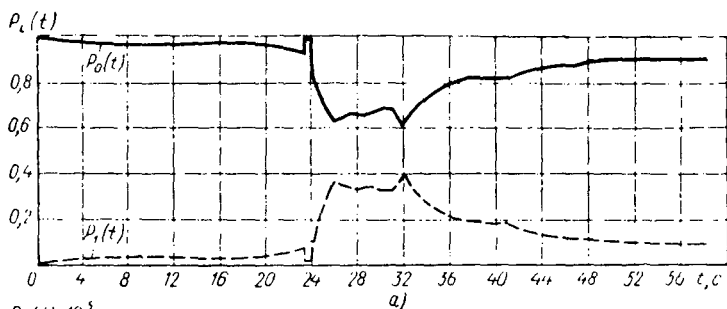


Рис. 2. 10. Графики переходных вероятностей двухконтурной СНС

Расчет должен завершиться сравнением полученных характеристик с заданным в ТЗ и в случае несоответствия — принятием корректирующих мер.

Во многих случаях представляет интерес решение задач анализа надежности СНС в замкнутой аналитической форме. В частности, это важно при постановке задачи оптимизации характеристики надежности СНС по одному из выбранных критериев.

Пример 2.2. Рассмотрим эволюцию состояний двухконтурной СНС (см. рис. 2.7), когда дифференциальный закон распределения времени адаптации аппроксимируется суперпозицией законов Эрланга нулевого, первого и второго порядков

$$f(u) = c_1 \mu_1 e^{-\mu_1 u} + c_2 \mu_2^2 u e^{-\mu_2 u} + c_3 \mu_3^3 u^2 e^{-\mu_3 u}. \quad (2.38)$$

Мгновенная интенсивность компенсируемых параметрических отказов — периодическая функция времени $v(t) = a(\sin \omega t + 1)$.

Относительно законов распределения других случайных потоков сохраняют силу сделанные в примере 2.1 допущения.

При этих условиях нарушается марковский принцип независимости будущего от прошлого при известном настоящем. С помощью приема, описанного в п. 2.4, сводим немарковский процесс развития состояний к марковскому. Расширенное пространство строится из 10 состояний (рис. 2.11). Эволюция состояний образует неоднородную цепь Маркова и описывается системой уравнений:

$$\begin{aligned}
 \dot{P}_0 &= -[\bar{v}(t) + \lambda + \eta] P_0(t) + \mu_1 P_5(t) + \mu_2 P_6(t) + \mu_3 P_3(t); \\
 \dot{P}_1 &= c_1 \bar{v}(t) P_0(t) - [\mu_2 + \lambda + \eta] P_3(t); \\
 \dot{P}_2 &= \lambda P_0(t) - [\bar{v}(t) + \eta] P_2(t); \\
 \dot{P}_3 &= -\eta P_3(t) + \bar{v}(t) P_2(t) + \lambda \left[P_1(t) + \sum_{j=5}^9 P_j(t) \right]; \\
 \dot{P}_4 &= \eta \sum_{i=0}^9 P_i(t), \quad i \neq 4; \\
 \dot{P}_5 &= c_2 \bar{v}(t) P_0(t) - [\mu_1 + \lambda + \eta] P_5(t); \\
 \dot{P}_6 &= \mu_2 P_1(t) - [\mu_2 + \lambda + \eta] P_6(t); \\
 \dot{P}_7 &= c_3 \bar{v}(t) P_0(t) - [\mu_3 + \lambda + \eta] P_7(t); \\
 \dot{P}_8 &= \mu_3 P_7(t) - [\mu_3 + \lambda + \eta] P_8(t); \\
 \dot{P}_9 &= \mu_3 P_8(t) - [\lambda + \eta + \mu_3] P_9(t).
 \end{aligned}
 \tag{2.39}$$

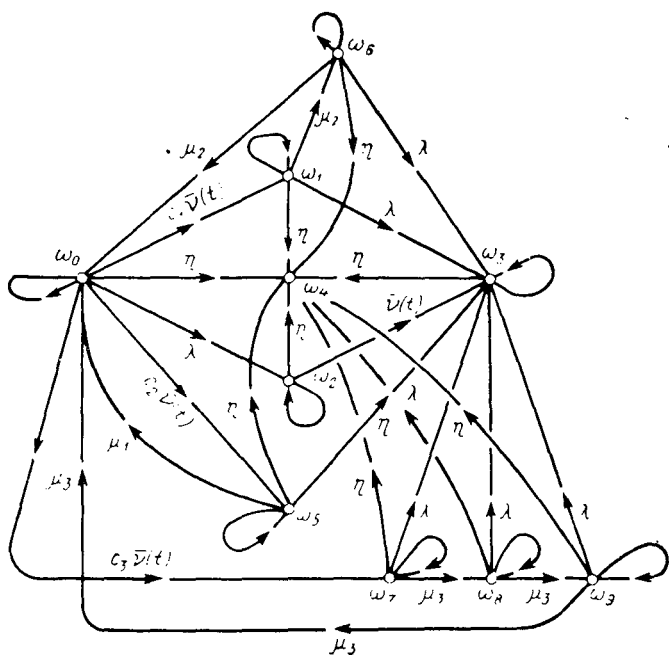


Рис. 2.11. Граф переходов двухконтурной СНС с расширенным фазовым пространством

Требуется получить аналитическое решение этой системы уравнений при начальных условиях:

$$P_0(0) = 1; P_i(0) = 0; i = 1, 2, \dots, 9. \quad (2.40)$$

Ниже приводятся результаты буквенного решения системы на ЭВМ «МИР-2» [50]. При решении был использован алгоритмический язык «АНАЛИТИК», разработанный в Институте кибернетики АН УССР под руководством В. М. Глушкова [10]. Решение дается здесь в точности в том виде, в котором оно выведено на печать:

$$P_0(X) = 1 + (-A - L - \Theta) \times X + 500\,000_{10} - 6 \times \\ \times (-V \times \Lambda - ((-A - L - \Theta) \times (\Lambda + L + \Theta)) + \\ + C_2 \times A \times M_1) \times X^{\uparrow 2};$$

$$P_1(X) = C_1 \times A \times X + 500\,000_{10} - 6 \times ((V - (A - L - \Theta)) \times \\ \times \Lambda \times C_1 \times A \times (M_2 + L + \Theta)) \times X^{\uparrow 2};$$

$$P_2(X) = L \times X + 500\,000_{10} - 6 \times ((-A - L - \Theta) \times \\ \times L - (L \times (A + \Theta))) \times X^{\uparrow 2};$$

$$P_3(X) = 500\,000_{10} - 6 \times (L \times A + (C_1 \times \Lambda + C_2 \times \Lambda + \\ + C_3 \times A) \times L) \times X^{\uparrow 2};$$

$$P_4(X) = 500\,000_{10} - 6 \times (C_1 \times A + L + C_2 \times \Lambda) \times \Theta \times X^{\uparrow 2};$$

$$P_5(X) = C_2 \times A \times X + 500\,000_{10} - 6 \times ((V - (A - L - \Theta)) \times \\ \times A \times C_2 - (C_2 \times \Lambda \times (M_1 + L + \Theta))) \times X^{\uparrow 2};$$

$$P_6(X) = 500\,000_{10} - 6 \times C_1 \times \Lambda \times M_2 \times X^{\uparrow 2};$$

$$P_7(X) = C_3 \times A \times X + 500\,000_{10} - 6 \times ((V - (A - L - \Theta)) \times \\ \times \Lambda \times C_3 - (C_3 \times A \times (M_3 + L + \Theta))) \times X^{\uparrow 2};$$

$$P_8(X) = 500\,000_{10} - 6 \times C_3 \times \Lambda \times M_3 \times X^{\uparrow 2};$$

$$P_9(X) = 0.$$

Решение системы (2.39) представлено первыми тремя членами разложения в ряд Тейлора¹.

Пример 2.3. Представляет интерес проследить за поведением переходных вероятностей СНС в том случае, когда мгновенная интенсивность параметрических отказов — периодическая функция времени.

Для удобства обратимся к уже рассмотренной СНС (см. рис. 2.7), но несколько изменим исходные данные, а именно; $\bar{v}(t) =$

¹ Для дешифровки полученного решения достаточно знать простой ключ:

$$a \rightarrow A \quad P_0 \rightarrow P_0 \quad c_1 \rightarrow C_1 \quad \mu_1 \rightarrow M_1$$

$$\omega \rightarrow V \quad P_1 \rightarrow P_1 \quad c_2 \rightarrow C_2 \quad \mu_2 \rightarrow M_2$$

$$t \rightarrow X \quad . \quad c_3 \rightarrow C_3 \quad \mu_3 \rightarrow M_3$$

$$\lambda \rightarrow L \quad \vdots$$

$$\eta \rightarrow \Theta \quad P_9 \rightarrow P_9$$

Символ $X^{\uparrow 2}$, как и в языке «АЛГОЛ-60», представляет собой X^2 ; $500\,000_{10} = 0,5$.

$= a(\sin \omega t + 1)$; $a = 0,502 \text{ с}^{-1}$; $\omega = 0,104 \text{ рад/с}$; $\mu = 1 \text{ с}^{-1}$; $\lambda = 15,3 \cdot 10^{-7} \text{ с}^{-1}$; $\eta = 43 \cdot 10^{-7} \text{ с}^{-1}$; $\tau = 130 \text{ с}$.

Требуется определить вероятность безотказной работы СНС за 130 с.

Решение

Система уравнений состояний будет иметь вид:

$$\left. \begin{aligned} \dot{P}_0 &= -[a(\sin \omega t + 1) + \lambda + \tau] P_0(t) + \mu P_1(t); \\ \dot{P}_1 &= [a(\sin \omega t + 1)] P_0(t) - (\mu + \lambda + \eta) P_1(t); \\ \dot{P}_2 &= \lambda P_0(t) - [a(\sin \omega t + 1) + \tau] P_2(t); \\ \dot{P}_3 &= \lambda P_1(t) + [a(\sin \omega t + 1)] P_2(t) - \eta P_3(t); \\ \dot{P}_4 &= \tau \sum_{i=0}^3 P_i(t). \end{aligned} \right\} \quad (2.41)$$

Начальные условия:

$$P_0(0) = 1; P_i(0) = 0; i = 1, \dots, 4. \quad (2.42)$$

Значения $v(t)$ и μ специально выбраны таким образом, чтобы на отдельных элементах траектории полета приведенная плотность потока параметрических отказов $\psi(t) > 1$.

На рис. 2.12 приводятся результаты решения системы (2.41) на ЦВМ. Переходные вероятности P_0 , P_1 и P_2 — периодические функции времени. Выделяются участки времени работы (в окрестности $t = 15 \text{ с}$

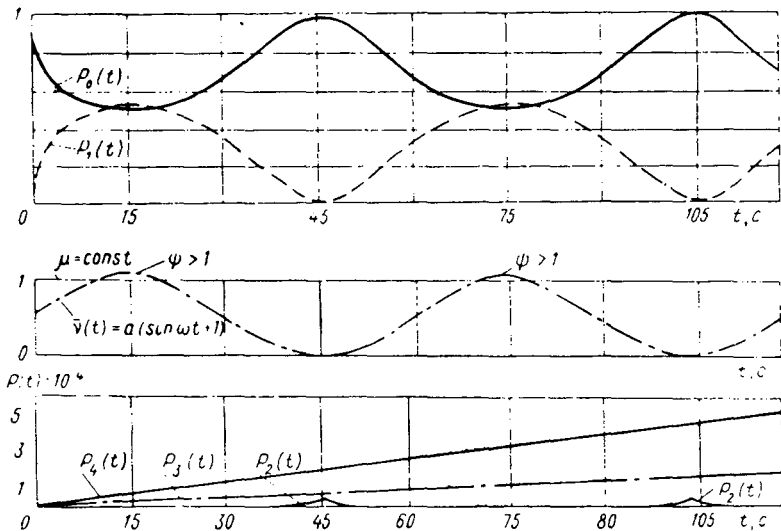


Рис. 2.12. Графики переходных вероятностей СНС для периодической интенсивности входящего потока

и $t=75$ с), на которых скорость самонастройки меньше темпа возникновения параметрических отказов. На этих участках наблюдается рассогласование между желаемым и действительным поведением СИС, а $P_1(t) > P_0(t)$. Вероятность попадания в поглощающий экран $P_4(t)$ — монотонно возрастающая функция, $P_2(t)$, $P_3(t)$ остаются пренебрежимо малыми на всем протяжении времени работы системы. Окончательно имеем $P_c(130 \text{ с}) = 0,999241$.

2.6. О СТРУКТУРЕ БОРТОВЫХ ЦИФРОВЫХ ВЫЧИСЛИТЕЛЬНЫХ МАШИН

Расширяющееся применение БЦВМ в контурах управления ЛА не является данью модному течению, а диктуется практической необходимостью: БЦВМ позволяют относительно просто реализовать сложные нелинейные корректирующие фильтры, алгоритмы самонастройки и осуществлять глубокий контроль подсистем ЛА [1, 2].

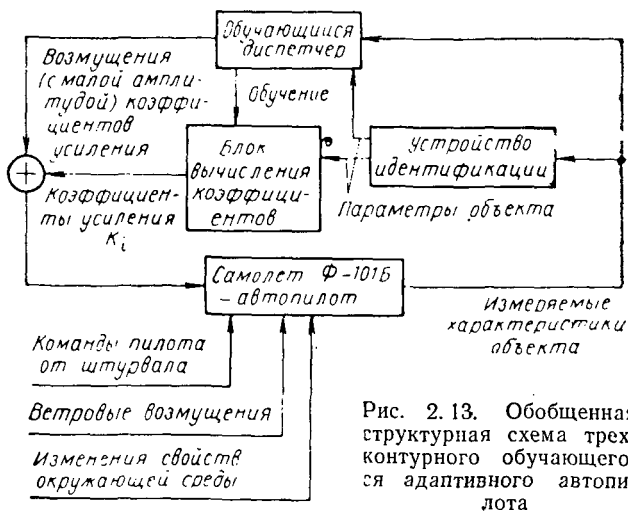


Рис. 2.13. Обобщенная структурная схема трехконтурного обучающегося адаптивного автопилота

При синтезе цифрового автопилота необходимо определить следующие параметры:

- время цикла выполнения программы;
- быстродействие, характеризующееся временем выполнения короткой операции;
- систему команд;
- разрядность чисел;
- объем запоминающего устройства.

Время цикла выполнения программы определяется максимально допустимым значением шага дискретизации времени для самого динамичного алгоритма управления, реализуемого с помощью БЦВМ.

Выбор типа БЦВМ определяется сложностью алгоритмов, скоростью протекания процессов, требуемой точностью реализации алгоритмов, допустимым весом и габаритами, надежностью, возможностью организации самоконтроля, а также контроля и диагностики других подсистем ЛА [2].

Измеряемые
характеристики объекта

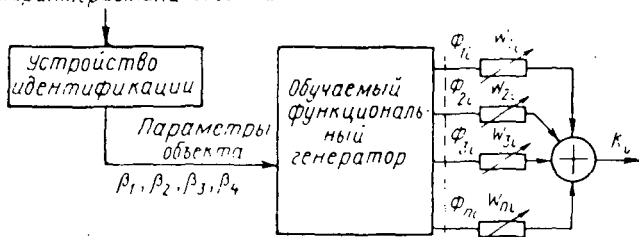


Рис. 2.14. Структурная схема контура самонастройки

Рассмотрим пример применения БЦВМ в обучающейся адаптивной системе управления полетом скоростного многорежимного самолета Ф-101Б [66]. Первоначально система была задумана как средство для повышения запаса продольной устойчивости, но разработанные методы оказались пригодными для решения разнообразных задач стабилизации и управления при различных функционалах, характеризующих качество управления (рис. 2.13).

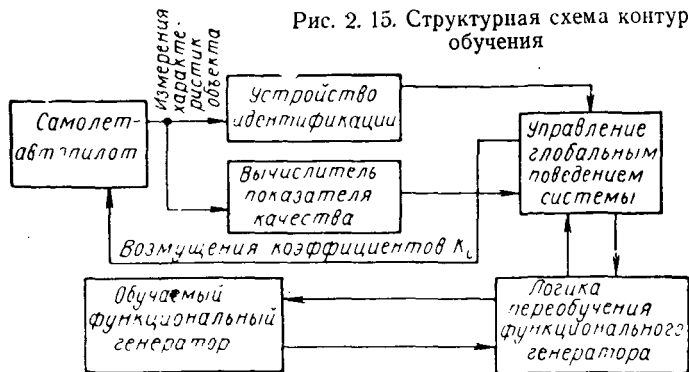
В этом автопилоте предусматриваются три контура: — внутренний (основной) контур «самолет — аналоговый автопилот», минимизирующий математическое ожидание интеграла от квадрата ошибки;

— быстродействующий контур самонастройки, реализованный в БЦВМ и включающий в себя функциональный преобразователь (рис. 2.14);

— «медленнодействующий» самообучающийся контур, который также воплощается в БЦВМ, управляет глобальным поведением системы и в случае необходимости осуществляет обучение функционального преобразователя (рис. 2.15).

Автопилот обладает способностью обучаться оптимальному и субоптимальному управлению в ранее не встречавшихся обстоятельствах. Улучшения в поведении, которые явились следствием приспособления к определенной обстановке, «вспоминаются», когда самолет вновь в нее попадает. Это происходит независимо от того, сколько времени истекло с тех пор, как система впервые столкнулась с этими обстоятельствами.

Рис. 2. 15. Структурная схема контура обучения



Таким образом, в трехконтурном автопилоте используется стратегия выборки информации о передаточных числах АП из ДЗУ БЦВМ и избирательного поощрения с целью поддержания относительно постоянных характеристик управления при существенных вариациях внутренней и внешней среды ЛА. При изменении параметров объекта контур самонастройки перестраивает передаточные числа аналогового автопилота. Самообучающийся контур корректирует те априорные допущения о свойствах объекта, которые оказались неточными.

Приведем некоторые подробности устройства и работы трехконтурного автопилота.

Основной контур. Критерий качества для синтеза оптимальной системы был взят в виде

$$I = M \left[\int_0^{\infty} (c_{ж} - j_y - 12,4\dot{\theta})^2 dt \right], \quad (2.43)$$

где $c_{ж} = (1 - e^{-3t})h_c$ — желаемый выход;
 h_c — ступенчатый управляющий входной сигнал;

j_y — отклонение от расчетного значения нормального ускорения;

$\dot{\theta}$ — угловая скорость тангажа.

Отметим, что из-за ошибок измерений и порывов ветра j_y и $\dot{\theta}$ в выражении (2.43) — случайные функции. Минимизация функционала (2.43) распадалась на две самостоятельные задачи. Первая заключалась в оптимальной оценке состояния объекта, основанной на измерениях величин j_y и $\dot{\theta}$. Во второй задаче требовалось найти оптимальное управление в предположении, что шум отсутствует. При таком подходе функционал качества не содержит случайных переменных.

В результате решения задачи минимизации был получен закон управления, в котором используются оптимальные оценки. В простейшем случае, когда уравнения движения линейные с постоянными коэффициентами, минимизация дает закон управления в виде

$$\delta_n = K_1 Z_\alpha \alpha + K_2 \dot{\theta} + K_3 \frac{W_{\text{п}}}{318} + K_4 c_{\text{ж}} + K_5 h_{\text{с}}, \quad (2.44)$$

где K_1, \dots, K_5 — коэффициенты усиления;

δ_n, α — углы отклонения от расчетных значений соответственно угла отклонения рулевого органа δ_n^* , угла атаки α^* ;

Z_α — производная подъемной силы по углу α ;

$W_{\text{п}}$ — скорость порыва ветра.

Для удобства математического моделирования было признано целесообразным использовать величины $Z_\alpha \alpha$, $\dot{\theta}$ и $W_{\text{п}}/318$, а не α , $\dot{\theta}$ и $W_{\text{п}}$. Состояние системы определялось с помощью фильтра Калмана с постоянными коэффициентами, который сглаживает измерения с шумом и дает на выходе α , $\dot{\theta}$ и $W_{\text{п}}$ (или, точнее, $Z_\alpha \alpha$, $\dot{\theta}$ и $W_{\text{п}}/318$). Коэффициенты фильтра устанавливались заранее и являлись оптимальными только для одного режима полета.

Для получения закона управления (2.44) и начальных значений коэффициентов усиления для всего диапазона изменений условий полета применялась теория оптимального управления, основанная на использовании метода наименьших квадратов (подробнее об этом см. [18]).

Контур адаптации состоит из устройства идентификации и обучаемого функционального преобразователя (см. рис. 2.14).

Устройство идентификации работает от измеряемых величин (скорости тангажа и нормального ускорения) и известных входов. Его входными величинами являются четыре обобщенных параметра $\beta_1, \beta_2, \beta_3, \beta_4$, характеризующие объект управления. Эти параметры связаны с производными $M_{\theta}, M_{\alpha}, M_{\delta_B}, M_{\dot{\alpha}}, Z_{\alpha}$ следующим образом:

$$\left. \begin{aligned} \beta_1 &= M_{\theta} + M_{\alpha}; & \beta_2 &= M_{\delta_B} + Z_{\delta_B} M_{\dot{\alpha}}; \\ \beta_3 &= \frac{M_{\alpha}}{Z_{\alpha}} + M_{\dot{\alpha}}; & \beta_4 &= Z_{\alpha}. \end{aligned} \right\} \quad (2.45)$$

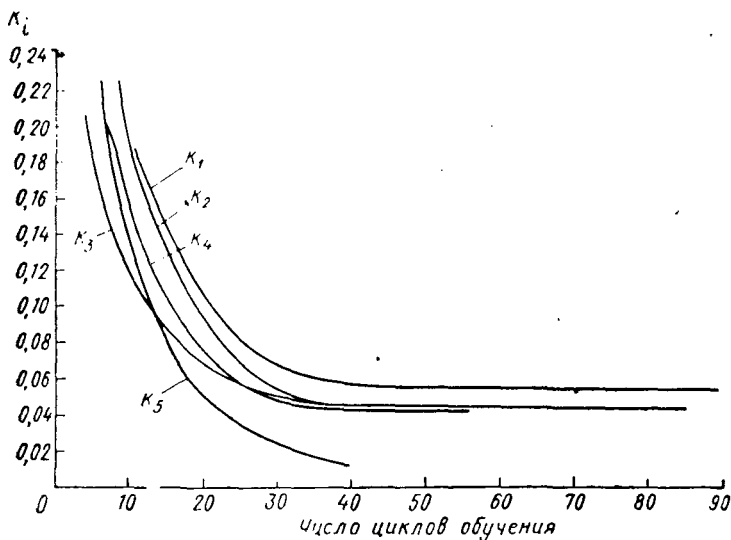


Рис. 2.16. К описанию процесса обучения для определения передаточных чисел автопилота

Задача устройства идентификации — настроить эти параметры так, чтобы поведение модели приближалось (в соответствии с выбранным критерием качества) к поведению ЛА. Уравнения устройства идентификации выведены с применением байесовского подхода.

Структура другого звена контура адаптации — обучаемого функционального преобразователя — иллюстрируется на рис. 2.14. Над множеством идентифицированных параметров объекта осуществляется фиксированное (не

зависящее от времени) преобразование, в результате которого получается другое множество величин, т. е. $\Phi = T(\beta)$. Полученные величины умножаются на веса, а получающиеся в результате произведения суммируются. Таким образом

$$K_i = \sum_{j=1}^n w_{ij} \Phi_j.$$

Для генерирования каждого коэффициента усиления обучаемый функциональный преобразователь должен располагать набором весовых коэффициентов. Первоначально преобразователь был обучен с помощью коэффициентов усиления, значения которых для 37 режимов полета были получены аналитически. Весь диапазон изменения условий полета квантовался на участки по скорости — 0,2 М и по высоте — 3 048 м. Для каждого из этих режимов были подсчитаны значения β_i и затем вычислены соответствующие коэффициенты усиления.

Диапазоны изменения величин β_i и рассчитанных коэффициентов K_1, \dots, K_5 указаны в табл. 2.3 и 2.4.

Система обучалась с применением метода наискорейшего спуска. Сходимость процесса обучения иллюстрируется графиками на рис. 2.16. Максимальная относительная ошибка в вычислении передаточных чисел дана здесь в зависимости от числа полных циклов обучения (каждый цикл соответствует определенному режиму полета Ф-101Б).

Таблица 2.3

Параметр объекта	Диапазон значений
β_1	от -3,1 до -0,369
β_2	от -58,9 до -2,79
β_3	от 3,99 до 71,6
β_4	от -2,02 до -0,217

Контур обучения (см. рис. 2.15). Задача контура — руководить глобальным поведением системы и определять возможно ли существенное улучшение качества ее работы. Для этого применяется градиентный поиск с использованием малых возмущений коэффициентов усиления. Когда приращение «улучшения» превысит некоторый заданный порог, производится переобучение функционального преобразователя таким образом, чтобы он выдавал «усовершенствованный» набор весовых коэффициентов для данного режима полета.

Таблица 2.4

Выходные величины функционального преобразователя	Желаемый диапазон значений
Коэффициент усиления в цепи обратной связи (по углу атаки) K_1	от $-1,92$ до $-20,7$
Коэффициент усиления в цепи обратной связи (по θ) K_2	от $1,11$ до $27,8$
Коэффициент усиления в цепи обратной связи от внешнего возмущения (скорости порыва ветра) K_3	от $-0,164$ до $-0,464$
Коэффициент усиления предварительного фильтра (от управляющего сигнала) K_4	от $-0,0685$ до $-1,87$
Коэффициент усиления разомкнутой цепи (от управляющего сигнала) K_5	от $-0,00663$ до $-0,146$

Критерий качества системы выражен функционалом

$$I = \frac{1}{2} \sum_{k=0}^{k'} (c_{и}(k) - c_{ж}(k))^2 e^{-\alpha(k'-k)},$$

где $c_{и}(k)$ — измеренное значение нормального ускорения ($+12,4 \theta$) в момент времени kT ; $c_{ж}(k)$ — желаемый выход. Экспоненциальный множитель используется для учета старения информации¹.

Задача заключается в выборе таких значений коэффициентов усиления автопилота, при которых минимизируется показатель качества.

Градиент функционала по передаточным числам в цепях обратной связи имеет вид

$$\nabla_k I = \sum_{k=0}^{k'} e^{-\alpha(k'-k)} (c_{и}(k) - c_{ж}(k)) \nabla_k c_{и}(k).$$

Градиент измеряемого выходного сигнала $\nabla_k c_{и}(k)$ нельзя получить непосредственно. Однако можно получить величину $\nabla_k \hat{c}(k)$, где $\hat{c}(k)$ — выходной сигнал модели устройства идентификации. Поэтому предполагалось, что

$$\nabla_k c_{и}(k) \approx \nabla_k \hat{c}(k).$$

¹ Устаревание данных всегда имеет место при конечном времени наблюдения (накопления) сигналов и вызывается изменением значений измеряемых параметров в процессе полета.

Поскольку известны и ошибка системы и приближенное значение градиента, для коррекции передаточных чисел в контуре обучения можно использовать метод наискорейшего спуска.

В системах управления космическими аппаратами длительного действия (2—3 года) и возможно в некоторых других конструкциях представляется целесообразной такая структура БЦВМ, при которой информация от многочисленных датчиков поступает сначала не в процессор, а в буферную память (БП), где записывается и ждет выборки по запросу процессора (рис. 2.17). Если

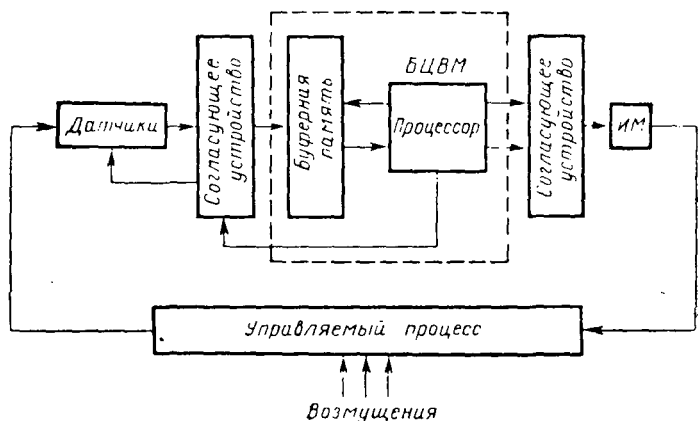


Рис. 2.17. Структурная схема САУ с БЦВМ:

ИМ — исполнительный механизм

все ячейки БП заняты, то вновь поступившая информация теряется. Из БП она может поступать в процессор, если в нем закончилось решение задачи. Эти задачи ответственны. Поэтому естественно требовать, чтобы вероятность потери информации не превышала заранее заданной малой величины. Эту величину вместе с требуемыми значениями характеристик надежности следует записывать в заданье на проектирование системы.

При известном темпе поступления и известном характере задач можно сформулировать такие требования к быстродействию БЦВМ и ее математическому обеспечению, которые исключали бы образование очереди. Однако, на любом этапе развития вычислительной техники мы сталкиваемся с противоречием между требуемым и

Пример 2.4. Структура проектируемой БЦВМ представлена на рис. 2.17.

Исходные данные.

1. Вероятность потери информации в результате заполнения БП не должна превышать $P_{\text{пот}}=10^{-3}$.

2. Математическое ожидание времени решения одной задачи в процессоре $t_{\text{обс}}=2$ с.

3. Плотность заявок на входе БП $\lambda=5$ с $^{-1}$.

Определить необходимый объем БП.

Решение.

1. Определяем

$$\psi \approx \lambda \bar{t}_{\text{обс}} = 2 \cdot 5 = 10.$$

2. По формуле (2.51) находим

$$n \approx \frac{\ln 10^{-3}}{\ln 10 - \ln 11} = 75.$$

На рис. 2.18 и 2.19 представлены графики зависимостей n от ψ при $P_{\text{пот}}=10^{-2}$ и $P_{\text{пот}}=10^{-3}$ соответственно. На графике рис. 2.19 обращает на себя внимание крутое возрастание объема БП при $\psi > 30$.

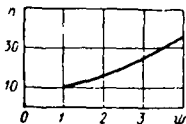


Рис. 2.18. График зависимости n от приведенной плотности потока заявок ψ при $P_{\text{пот}}=10^{-2}$

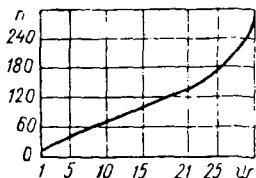


Рис. 2.19. График зависимости n от ψ при $P_{\text{пот}}=10^{-3}$

В заключение заметим, что рассмотрение БЦВМ и вычислительных систем с разделением времени с позиций теории массового обслуживания как многофазных СМО — перспективное направление исследований (см. например, работы [3, 48]). Оно находится на одной из своих начальных стадий и предстоит проделать большую работу, раньше чем результаты этого направления можно будет рекомендовать для структурного синтеза БЦВМ.

Подведем итоги изложенному.

1. Прежде чем делать выбор в пользу адаптивной системы, необходимо удостовериться в том, что она обладает явными преимуществами по сравнению с хорошо спроектированной и тщательно отработанной в ходе испытаний более простой стационарной системой. Таким образом удастся избежать неоправданного обстоятельствами летной эксплуатации усложнения оборудования.

Нужно иметь в виду, что контур адаптации может быть «введен в заблуждение» как некоторыми из своих собственных неисправностей (преимущественно типа «КЗ»), так и внешними факторами: флуктуацией питающего напряжения, турбулентностью атмосферы, упругими колебаниями конструкции ЛА. Следствием этих явлений может стать ложная адаптация.

2. Адаптация как метод обеспечения надежности систем оказывается целесообразной при неполной начальной информации об объекте, возмущающих воздействиях и начальных условиях. Ее применение эффективно в сочетании со структурной избыточностью на микро- и макроуровнях.

3. Интегро-дифференциальные уравнения вида (2.10) и уравнения Колмогорова являются адекватными моделями надежности адаптивных систем. Они позволили связать в единый ансамбль такие характеристики адаптивной системы, как скорость адаптации, суммарные интенсивности отказов в основном контуре и контуре самонастройки со свойствами окружения и требованиями к качеству регулирования. В результате их решения получены характеристики безотказности адаптивной системы: вероятность безотказной работы и математическое ожидание времени до первого попадания в поглощающий экран.

4. Изучение переходных вероятностей дает возможность проследить эффект запаздывания адаптации относительно изменений среды, наблюдаемый при работе адаптивных систем с обратной связью по критерию качества.

Таким образом, анализ решений системы интегро-дифференциальных уравнений вида (2.10) и уравнений Колмогорова позволяют наблюдать не только результат адаптации (достижение цели в меняющемся окружении), но и некоторые черты ее процесса.

НЕКОТОРЫЕ ЗАДАЧИ УПРАВЛЕНИЯ ДВИЖЕНИЕМ НА ВОЗДУШНЫХ ЛИНИЯХ

3.1. УПРАВЛЕНИЕ ПОСАДКОЙ ГРУПП САМОЛЕТОВ

Интенсификация движения на воздушных линиях и обеспечение безопасности полета требуют решения комплекса сложных организационных и технических задач.

Эти задачи распадаются на три части, связанные со взлетом, полетом по трассе и посадкой самолетов. Одной из наиболее ответственных задач является управление посадкой групп самолетов. Если мы и рассмотрим.

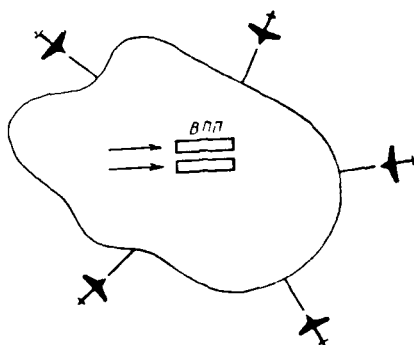


Рис. 3.1. Схема аэродромной зоны

Определим начальные условия следующим образом. Аэропорт окружен некоторой областью, называемой аэродромной зоной (рис. 3.1), в пределах которой осуществляется регулирование движения самолетов. Эта зона может быть любой

формы, а ее граница в зависимости от ряда обстоятельств меняется от нескольких десятков до нескольких сотен километров. Предполагается, что вне аэродромной зоны самолеты летят к аэродрому по радиальным направлениям. Время подхода к границе области и направление подхода случайны.

Обычно сначала бывает довольно трудно принять гипотезу о случайных моментах подхода к границе аэро-

ромной зоны, поскольку полеты гражданских самолетов совершаются по расписанию. Однако, исследование действительного движения показывает, что разница между ожидаемым временем прибытия и прибытием может быть величиной того же порядка, что и средний интервал между посадками самолетов. Кроме того, для простоты будем предполагать, если не оговорено противное, что на аэродроме совершаются только посадки.

Собственно посадкой называется движение самолета с высоты 400—500 м до приземления и полной остановки. Это движение может совершаться по жесткой пространственной траектории, называемой глиссадой, или по нежесткой траектории. В первом случае требуемая траектория движения (глиссада) задается с помощью размещенных на земле устройств, а отклонения центра масс самолета от глиссады измеряются бортовыми приборами. Во втором случае информация о положении самолета по отношению к точке приземления получается с помощью бортовых приборов.

Различают [11] заход на посадку, когда самолет совершает движение по глиссаде до высоты 20—30 м, и посадку, представляющую маневр в вертикальной плоскости, заканчивающийся приземлением самолета и движением по взлетно-посадочной полосе (ВПП) до полной остановки. В настоящее время общепризнано, что из всех этапов полета наиболее трудным является управление самолетом при заходе на посадку и посадке.

Для облегчения управления самолетом при посадке применяются различные системы захода на посадку, облегчающие ориентировку и позволяющие удерживать самолет на глиссаде. Поскольку на посадочных режимах приходится пользоваться не только приборами, дающими информацию о положении самолета на глиссаде, но и обычными приборами (вариометр, высотомер, компас, авиагоризонт и др.), то ясно, что ручное управление посадкой крайне затруднительно. Управление на посадочных режимах при наличии информации о положении самолета по отношению к заданной глиссаде может быть автоматизировано. Для этого указанная информация должна быть подана в автопилот¹. Последний получает управляющие сигналы от измерителей координат траектории посадки. При полуавтоматическом управлении

¹ Описание захода на посадку одиночного самолета с незначительными изменениями заимствовано из гл. XI книги [11].

удержание самолета на траектории, осуществляет легчайше на основе информации от приборов, характеризующих отклонение самолета от заданной траектории посадки.

Управление самолетом при заходе на посадку является частным случаем управления на заданной траектории полета; при этом в отличие от обычных маршрутных полетов требования к точности стабилизации на глиссаде посадки являются весьма высокими.

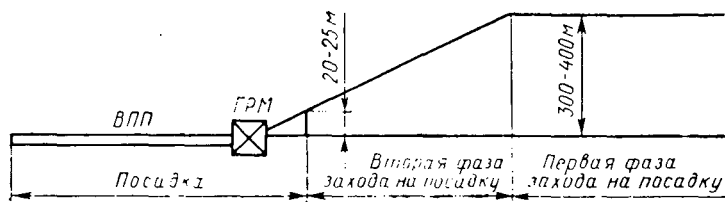


Рис. 3. 2. Схема захода самолета на посадку

Поскольку при заходе на посадку важно не только обеспечить движение самолета по глиссаде, но поддерживать также определенную скорость полета, превышающую на 20—30% критическую скорость самолета, то управление должно осуществляться как через управляющие поверхности, так и через тягу двигателей.

Во всех известных системах автоматического и полупод автоматического управления при заходе на посадку для задания посадочной глиссады используются радиотехнические устройства в сочетании с бортовым приборным оборудованием самолетов.

Заход на посадку состоит из двух фаз. В первой фазе самолет, предварительно совершив необходимые маневры, заходит на линию, совпадающую с центральной линией ВПП, имея постоянную высоту полета (рис. 3. 2). Эта линия заканчивается в точке пересечения с линией глиссады, направленной под углом 3—4° к горизонту. С этого момента начинается вторая фаза захода на посадку, в которой самолет снижается по глиссаде до высоты 20—25 м.

Радиотехническая система захода на посадку по приборам состоит из аэродромного и самолетного оборудования. Аэродромное оборудование системы состоит из глиссадного радиомаяка ГРМ, курсового радиомаяка КРМ, маркерных радиомаяков (ближнего маркера БМ, среднего маркера СМ и дальнего маркера ДМ), обзорного радиолокатора, радиосвязного оборудования и других подсистем (рис. 3. 3). Самолетное оборудование системы состоит из курсового приемника, глиссадного приемника, дальномера и указательного прибора. Курсовой и глиссадный радиомаяки излучают по два сигнала с перекрывающимися диаграммами направленности, и образуется равносигнальная зона, представляющая собой глиссаду.

Для приема сигналов курсового и глассадного маяков на самолете применяется одна ненаправленная антенна. Принятые антенной сигналы в виде напряжений разделяются полосовыми фильтрами выпрямляются и вычитаются, в результате получается разность напряжений постоянного тока, которая подается в нуль-прибор. Вертикальная и горизонтальная стрелки отклоняются на полную шкалу при уходе самолета от посадочной глассады на 2—3°. В пределах этого диапазона шкалы равномерны.

Самолет начинает снижение при прохождении дальнего маркера находясь на высоте 300—400 м. При полете ближнего маркера самолет должен находиться на высоте 16—25 м.

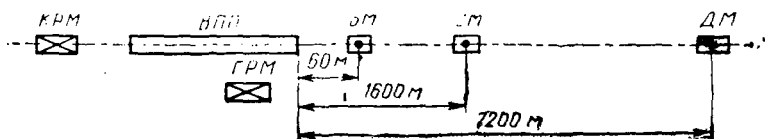


Рис. 3.3. Схема аэродромного оборудования посадки

Преимуществом радиотехнической системы является то, что она допускает сравнительно простую связь с автопилотом¹. Выходные сигналы курсового и глассадного приемников этой системы, представляющие сигналы постоянного тока, можно подать после фильтрации и усиления в соответствующие каналы управления автопилота. Если сигнал курсового приемника подать в автопилот, то самолет будет совершать разворот, скорость которого пропорциональна значению этого сигнала.

Вывод уравнений продольного и бокового движения самолета при заходе на посадку и посадке, анализ их решений, формирование законов регулирования тяги и управления рулями входят в задачи теории автоматического управления полетом.

Нас же будет интересовать не инфинитезимальное поведение самолета при посадке, а временные и отчасти качественные характеристики посадки группы самолетов. Тогда входящий поток требующих посадки самолетов, вспомогательные навигационные средства и ВПП можно рассматривать с позиций теории массового обслуживания как разомкнутую многоканальную СМО.

При выводе формул (1.1) — (1.12), характеризующих эффективность систем этого класса, основным было предположение о пуассоновском характере входящего потока заявок и экспоненциальном законе распределения вре-

¹ В настоящее время в широких масштабах проводится стандартизация характеристик курсоглассадных радиомаячных систем на основе Международной конвенции по вопросам гражданской авиации. Осуществление этой меры является одним из важнейших условий снижения минимумов погоды для посадки самолетов, повышения регулярности полетов на трассах и обеспечения безопасности движения в аэродромной зоне.

Если ВПП свободна, то первый самолет не задерживается и входит на глиссаду, а второму дается предписание ждать разрешения на посадку (летать по кругу) в аэродромной зоне. При высокой интенсивности движения может образоваться очередь ожидающих разрешения на посадку самолетов.

Функция плотности случайных интервалов между самолетами на глиссаде получается из следующих соображений. Посадки самолетов невозможны, если интервалы между ними меньше минимально безопасных, т. е. $t < \Delta$. Вероятность того, что самолеты поступают на глиссаду с интервалами $t > \Delta$, равна частоте их появления $1 - \lambda\Delta$, умноженной на вероятность того, что случайный интервал больше $t - \Delta$, т. е.

$$(1 - \lambda\Delta) P \{T > t - \Delta\} = (1 - \lambda\Delta) e^{-\lambda(t - \Delta)}. \quad (3.1)$$

Плотность распределения этих интервалов равна

$$f(t, \Delta | T > t - \Delta) = \lambda(1 - \lambda\Delta) e^{-\lambda(t - \Delta)}. \quad (3.2)$$

С учетом того, что часть самолетов поступает из очереди на глиссаду с интервалами Δ , общее выражение для плотности интервалов между посадками самолетов имеет вид

$$f(t, \Delta) = \lambda(1 - \lambda\Delta) H(t - \Delta) e^{-\lambda(t - \Delta)} + \lambda\Delta\delta(t - \Delta), \quad (3.3)$$

где $H(t - \Delta)$ — ступенчатая функция Хевисайда;

$\delta(t - \Delta)$ — дельта-функция Дирака.

Функция (3.3) является плотностью смешанного (дискретного и смещенного экспоненциального) распределения. Эта функция, как показали многочисленные исследования, хорошо согласуется с фактическим распределением случайных интервалов при высокой интенсивности движения.

Математическое ожидание и дисперсия длительности интервалов при регулировании входящего потока равны соответственно:

$$\bar{T} = \lambda\Delta\Delta + \frac{1 - (\lambda\Delta)^2}{\lambda}; \quad (3.4)$$

$$D = [1 - (\lambda\Delta)^2] \left[\frac{2}{\lambda^2} - \Delta^2 - \frac{1}{\lambda^2} (1 - (\lambda\Delta)^2) \right]. \quad (3.5)$$

При возрастании интенсивности движения или увеличении интервала безопасности пределы выражений (3.4) и (3.5) при $\lambda\Delta \rightarrow 1$ равны:

$$\bar{T} = \Delta; \quad D = 0, \quad (3.6)$$

т. е. имеет место детерминированный поток.

При малых интервалах безопасности или уменьшении интенсивности входящего потока, когда регулирование осуществляется редко, пределы выражений (3.4) и (3.5) при $\lambda\Delta \rightarrow 0$ равны:

$$\bar{T} = \frac{1}{\lambda}, \quad D = \frac{1}{\lambda^2}, \quad (3.7)$$

т. е. получается стационарный пуассоновский поток.

Из изложенного должно стать ясным, что лишь в частных случаях входящий поток требований на радиотехническую систему посадки обладает свойствами стационарности, ординарности и отсутствия последствия. Тем не менее, изучение поведения СМО с пуассоновским входящим потоком и экспоненциальным распределением времени обслуживания представляет большой интерес. Он связан с тремя обстоятельствами, отмеченными в п. 1.1.

После этих предварительных замечаний о структуре входящего потока рассмотрим примеры применения методов теории массового обслуживания к организации и управлению посадкой самолетов.

Пример 3.1. Поток прибывающих в аэропорт самолетов случайный пуассоновский с интенсивностью $\lambda=27$ прибытий в час. Случайный интервал времени, в течение которого ВПП занята, распределен по экспоненциальному закону. Среднее время обслуживания $\bar{t}_{обс} = 2$ мин.

Определить:

- 1) минимальное число ВПП, которое нужно предусмотреть на аэродроме, чтобы вероятность ожидания в воздухе для прибывающего самолета была бы меньше 0,1;
- 2) среднее время ожидания одним самолетом разрешения на посадку $\bar{t}_{ож}$;
- 3) среднюю длину очереди (среднее число самолетов, ожидающих в аэродромной зоне разрешения на посадку) \bar{n} ;
- 4) среднее число свободных ВПП N_0 ;
- 5) среднее число занятых ВПП N_z ;
- 6) коэффициент занятости K_z ;
- 7) коэффициент простоя $K_{п}$.

Решение

Систему посадки рассматриваем как разомкнутую СМО с ожиданием. Интересующие нас показатели эффективности СМО в стационарном режиме определяем по формулам (1.1)–(1.12). Ответ на первый вопрос ищем, применяя формулы (1.1), (1.2), (1.4) и алгоритм перебора.

Положим число ВПП равным 2, 3, 4. Определим:

$$\pi(n=2) = \frac{\alpha^n}{(n-1)!(n-\alpha)} P_0 = 0,278; \quad \pi(n=3) = 0,07;$$

$$\pi(n=4) = 0,014.$$

Значение $n=3$ является наименьшим, при котором выполняется неравенство $\pi(n) \leq 0,1$. Учитывая, что при фиксированном α π является убывающей функцией n , мы не ищем других значений n .

Среднее время ожидания определяем по формуле (1.5):

$$\bar{t}_{ож} = \frac{\pi \bar{t}_{обс}}{n-\alpha} = \frac{0,07 \frac{2}{60}}{3 - \frac{27}{30}} = 0,001 \text{ ч} = 3,6 \text{ с.}$$

Число самолетов, ожидающих разрешения на посадку:

$$\bar{v} = \lambda \bar{t}_{ож} = 27 \cdot 0,001 = 0,027, \text{ т. е. в среднем } 0 \text{ самолетов.}$$

Среднее число свободных ВПП:

$$N_0 = \sum_{k=0}^{n-1} \frac{n-k}{k!} \alpha^k P_0 = 2,098 \approx 2.$$

Среднее число занятых ВПП:

$$N_3 = n - N_0 = 1.$$

Коэффициент занятости ВПП:

$$K_3 = \frac{N_3}{n} = 0,33.$$

Коэффициент простоя ВПП:

$$K_{II} = \frac{N_0}{n} = 0,67.$$

Таким образом, для обеспечения заданной вероятности ожидания при фиксированных λ и $\bar{t}_{обс}$ требуется три ВПП, причем они используются только в течение 1/3 времени, а 2/3 времени простаивают.

Окончательное суждение о целесообразности такой организации системы посадки можно вынести, применяя экономические критерии. Выбор оптимального варианта структуры СМО, минимизирующего суммарный ущерб,

рассмотрен в задаче о дозаправке самолета горючим в полете (см. п. 3. 2).

Система массового обслуживания с ожиданием и приоритетом в обслуживании. До сих пор поток самолетов, ожидающих разрешения на посадку, рассматривался нами как однородный. Это, конечно, идеализация реальных обстоятельств, в которых совершается посадка самолетов. В первую очередь, разрешение дастся самолетам при наличии в них неисправностей или при других условиях, оговоренных в инструкциях аэродромной диспетчерской службы.

На аэродромах с одной ВПП самолетам, идущим на посадку, полоса предоставляется в первую очередь, т. е. они пользуются приоритетом по сравнению с другими самолетами, использующими полосу для взлета.

Профессором Л. А. Овчаровым [36] рассмотрена следующая задача.

Пример 3. 2. Изучается функционирование большого аэропорта с одной ВПП. В среднем за сутки взлетает 240 самолетов и столько же садится. При посадке самолет занимает ВПП в среднем 3 мин, а при взлете — 1,5 мин. Определить характеристики работы аэропорта в установившемся режиме.

Решение

Аэропорт рассматривается как одноканальная СМО с приоритетом. Самолету разрешается взлет в том случае, когда нет самолетов, идущих на посадку. Поток самолетов, идущих на посадку, можно рассматривать как поток требований, обладающих приоритетом в использовании ВПП. Характеристики такой СМО с приоритетом следующие:

$$\lambda_1 = \lambda_2 = 240 \text{ 1/сут} = \frac{1}{6} \text{ мин}^{-1}.$$

Математическое ожидание времени обслуживания:

$$\bar{t}_{\text{обс}_1} = \frac{1}{\mu_1} = 3 \text{ мин}, \quad \mu_1 = \frac{1}{3} \text{ мин}^{-1};$$

$$\bar{t}_{\text{обс}_2} = \frac{1}{\mu_2} = 1,5 \text{ мин}, \quad \mu_2 = \frac{2}{3} \text{ мин}^{-1}.$$

$$\text{Следовательно: } \alpha_1 = \frac{\lambda_1}{\mu_1} = 0,5; \quad \alpha_2 = \frac{\lambda_2}{\mu_2} = 0,25.$$

Условие существования стационарного режима $\alpha_1 + \alpha_2 < 1$ выполняется.

Среднее число самолетов, ожидающих в воздухе пока освободится ВПП,

$$\bar{v}_1 = \frac{(a_1)^2}{1 - a_1} = \frac{(0,5)^2}{1 - 0,5} = 0,51.$$

Среднее время ожидания разрешения на посадку:

$$\bar{t}_{ож_1} = \frac{\bar{v}_1}{\lambda_1} = 0,5 \cdot 6 = 3 \text{ мин.}$$

Среднее время, затрачиваемое самолетом на посадку,

$$\bar{t}_{пр_1} = \bar{t}_{ож_1} + \bar{t}_{обс_1} = 3 + 3 = 6 \text{ мин.}$$

Среднее время ожидания разрешения на взлет для самолета, находящегося на аэродроме:

$$\bar{t}_{ож_2} = \frac{1}{\mu_2} \frac{\mu_2 \frac{\alpha_1}{1 - \alpha_1} + \alpha}{1 - \alpha} = \frac{3}{2} \frac{2 \frac{0,5}{0,5} + 0,75}{0,25} = 16,5 \text{ мин.}$$

Среднее число самолетов, ожидающих на аэродроме разрешения на взлет:

$$\bar{v}_2 = \frac{\alpha_1}{(1 - 2\alpha_1)(1 - \alpha_1)} = \frac{1}{6} 16,5 \approx 3.$$

Среднее время, проходящее от момента готовности самолета к взлету с аэродрома до момента осуществления взлета:

$$\bar{t}_{пр_2} = \bar{t}_{ож_2} + \bar{t}_{обс_2} = 16,5 + 1,5 = 18 \text{ мин.}$$

Среднее число самолетов, готовых к взлету, но находящихся на аэродроме:

$$\bar{v}_2' = \lambda_2 \bar{t}_{пр_2} = \frac{1}{6} 18 = 3.$$

Из рассмотрения характеристик СМО видно, что несмотря на довольно большое число взлетопосадок за сутки (480) можно считать, что аэропорт будет функционировать «нормально»: самолетам, идущим на посадку, не придется долго ждать в воздухе разрешения на посадку.

Заметим, что при реальном функционировании аэропорта самолеты, идущие на посадку, не имеют абсолютного приоритета: если взлетающий самолет уже бежит по ВПП, то его взлет не прекращается. Однако, приближение не вносит сколько-нибудь заметных изменений в полученное решение.

¹ Если величина \bar{v}_1 используется в последующих вычислениях, то она не округляется до ближайшего целого.

3.2. УПРАВЛЕНИЕ ЗАПРАВКОЙ САМОЛЕТОВ В ВОЗДУХЕ

Вопросам увеличения дальности полета самолета и способности при минимальном запасе топлива доставлять грузы на предельные расстояния по-прежнему уделяется большое внимание. Одним из эффективных средств повышения дальности полета и полетной нагрузки считают системы заправки самолетов в воздухе (см., например, [47]).

В настоящее время системами заправки оснащаются не только бомбардировщики и военно-транспортные самолеты, но и истребители. В качестве заправщиков используются как специализированные, так и боевые самолеты, снабженные подвесным заправочным оборудованием. Специализированные заправщики применяются для дозаправки самолетов ВВС Англии, США, Франции и других стран.

Создание систем, обеспечивающих заправку в полете, идет несколькими путями и диктуется соображениями, связанными с выбором оптимальных параметров для различных целевых назначений.

Основные тенденции в развитии систем заправки таковы:

- расширение диапазона скоростей и высот заправки;
- сокращение времени обнаружения, стыковки и перелива;
- снижение веса системы;
- обеспечение автономности энергопитания;
- снижение нагрузок на топливоприемники заправляемого самолета;
- унификация заправочного оборудования.

Ведутся работы по автоматизации контактирования; при этом используется опыт наведения ракет на цель и стыковки космических аппаратов на орбите. Сколько-нибудь детальное рассмотрение применяемых за рубежом систем заправки и организации их эксплуатации выходит за рамки этой главы. Заинтересованного читателя отсылаем к работе [47].

Если пренебречь инфинитезимальным поведением системы заправки и учитывать лишь временные и отчасти качественные характеристики процесса заправки, то такая система с позиций теории массового обслуживания может быть рассмотрена как СМО смешанного типа.

Приведем иллюстративный пример, из которого становится ясным, какие характеристики системы заправки в полете могут быть определены при подходе к этой проблеме с позиций исследования операций и какие исходные данные являются для этого необходимыми и достаточными.

Пример 3.3. В районе дозаправки постоянно дежурит n самолетов-дозаправщиков. Если дозаправка началась, то она осуществляется до конца и длится в среднем 10 мин. Если все дозаправщики заняты, то самолет нуждающийся в дозаправке, некоторое время может ожидать (совершать полет по кругу в районе дозаправки). Среднее время ожидания — 20 мин. Если самолет так и не дожидается дозаправки, он садится на запасной аэродром. В среднем в район дозаправки за час прибывает 24 самолета. Число самолетов, ожидающих дозаправки, ничем не ограничено.

Определить оптимальное число самолетов-дозаправщиков, при котором минимизируется платежная функция

$$C(n) = c_1 \lambda \bar{t}_{ож} + c_2 K_{пн} + c_3 K_{zn}, \quad (3.8)$$

где $C(n)$ — суммарные потери в системе дозаправки горючим в воздухе за единицу времени (1 ч) ее работы;

c_1 — ущерб от ожидания одним самолетом дозаправки в воздухе за единицу времени;

c_2 — ущерб от простоя одного дозаправщика за единицу времени;

c_3 — затраты на эксплуатацию одного дозаправщика за единицу времени.

Принимаем: $c_1 = 50$ единиц стоимости; $c_2 = 25$ единиц стоимости; $c_3 = 25$ единиц стоимости.

Решение. По формулам (1.13)—(1.19) с помощью табл. 5 и 6, приведенных в книге [35], определяем характеристики гипотетической СМО.

1. Вычисляем параметры α и β :

$$\alpha = 24 \frac{10}{60} = 4,$$

$$\beta = \frac{10}{.60} \frac{60}{20} = 0,5.$$

2. На основе накопленного опыта в организации систем дозаправки ориентировочно определяем область допустимых значений числа каналов обслуживания. Пусть $n=1, 2, 3, 4, 5$. Для этих значений n определяем характеристики СМО и результаты сводим в табл. 3.1.

3. Определяем значения платежной функции для $n=1, \dots, 5$ и сводим результаты в табл. 3.2.

Из приведенных значений функции $C(n)$ следует, что $\min C(n)$ имеет место при $n=4$, т. е. искомое $n_{opt} = 4$.

На основании изложенного можно отметить, что некоторые задачи организации и управления воздушным

Таблица 3.1

n	λ	α	β	ρ_0	N_3	\bar{v}	$P_{0\text{ис}}$	$P_{0\text{гк}}$	K_3	$t_{\text{пр. ч}}$	$\bar{t}_{\text{ож. ч}}$	$K_{\text{п}}$
1	24	4	0,5	0,0027	0,9973	5,57	0,248	0,71	0,9973	0,00045	0,232	0
2	24	4	0,5	0,0073	1,956	4,05	0,49	0,51	0,977	0,00383	0,168	0
3	24	4	0,5	0,0121	2,8	2,4	0,7	0,3	0,927	0,0113	0,1	0,073
4	24	4	0,5	0,0151	3,3	1,17	0,83	0,17	0,845	0,0305	0,05	0,155
5	24	4	0,5	0,0171	3,7	0,64	0,92	0,08	0,737	0,0593	0,03	0,263

Таблица 3.2

n	$c_1 \lambda \bar{t}_{\text{ож}}$	$c_2 K_{\text{п}} n$	$c_3 K_3 n$	$C(n)$
1	278,4	0,075	24,9	303,3
2	201,5	1,1	49	251,6
3	120	5	70	195
4	60	17,5	82,5	160
5	36	42,5	92,5	171

движением приводят к необходимости изучения явлений ожидания и допускают формулировку на языке теории массового обслуживания.

Выбор наилучшей стратегии в организационных задачах, формулируемых в терминах теории массового обслуживания, подразумевает минимизацию суммарного ущерба, связанного с очередью заявок и простоем каналов.

ПЛАНИРОВАНИЕ ОБСЛУЖИВАНИЯ СИСТЕМ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

4.1. ПЛАНИРОВАНИЕ ЗАМЕНЫ ОТДЕЛЬНЫХ ЭЛЕМЕНТОВ СИСТЕМЫ ПРИ МГНОВЕННОЙ ИНДИКАЦИИ ОТКАЗОВ

В процессе работы САУ возникает необходимость в периодической замене отдельных наиболее важных и наименее надежных элементов.

Рассмотрим элемент с функцией распределения времени безотказной работы $F(t) = P(X < t)$ и монотонно возрастающей интенсивностью отказов ($F'(t) = f(t)$).

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \quad (4.1)$$

Здесь X — случайная величина времени от момента возвращения элемента в исходное состояние до отказа. Для элемента назначается профилактическое обслуживание (контроль) через случайное время Y , распределенное по закону

$$G(t) = P(Y < t).$$

Выполняется профилактическое обслуживание через $\min(X, Y)$ единиц времени. Другими словами, если после предыдущего возвращения в исходное начальное состояние (момент t_2 на рис. 4.1) элемент безотказно работал время Y (в этом случае $X > Y$), то профилактическое обслуживание проводится сразу же по истечении времени Y , причем время проведения профилактики предполагается случайным. Обозначим его символом V ; функция распределения величины V $F_V(t)$ произвольна, а математическое ожидание $M[V] = T_{11} < \infty$.

Если же после предыдущего возвращения в исходное состояние (после моментов времени $t=0$ и $t=t_4$) элемент отказал до истечения срока, равного $Y=y_1$ и $Y=y_3$ и т. д.

(см. рис. 4.1), т. е. при $X < Y$, то непосредственно после отказа происходит его замена. Время, затрачиваемое на замену, предполагается случайным и обозначается величиной U , имеющей произвольную функцию распределения $F_U(t)$ и математическое ожидание $M[U] = T_3 < \infty$. Будем считать, что $T_{II} < T_3$.

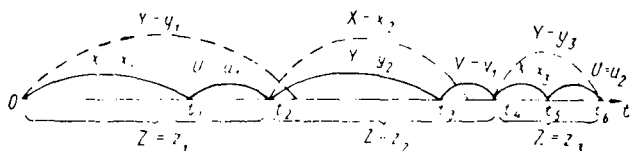


Рис. 4.1. К принципу осуществления замен через случайное время

Задача заключается в отыскании такого периода профилактического обслуживания, при котором элемент обладает максимальной вероятностью нахождения в исправном состоянии в произвольно выбранный момент t (максимальная готовность) и максимальной вероятностью исправного состояния на отрезке времени $t, t+x$ (максимальная надежность за время работы в воздухе x). Обозначим $p(x, t)$ вероятность того, что элемент окажется исправным в произвольный момент t и безотказно проработает в воздухе в течение времени $t+x$.

Время между очередными моментами возвращения элемента в исходное состояние обозначим через Z . Тогда из рис. 4.1 ясно, что

$$Z = \begin{cases} Y \dot{-} V, & \text{если } X > Y; \\ X \dot{-} U, & \text{если } X < Y. \end{cases} \quad (4.2)$$

Последовательные реализации случайной величины Z обозначим Z_k .

Величины Z_k образуют последовательность независимых, одинаково распределенных случайных величин, т. е. образуют процесс восстановления с $\mu = M[Z]$ и функцией восстановления $H(t) = M[N_t]$, где N_t — максимальное значение n , при котором $S_n \leq t$, где

$$S_n = Z_1 + Z_2 + \dots + Z_n.$$

Если случайные величины X и Y обладают плотностью, то $dH(t)$ можно рассматривать, как вероятность

того, что возвращение элемента в исходное состояние (восстановление) завершится в момент t , так как вероятность такого восстановления на малом отрезке Δt равна $H(t+\Delta t) - H(t) + o(\Delta t)$. Математическое ожидание числа восстановлений элемента на малом отрезке времени Δt равно

$$M[\Delta N_t] = p_1(\Delta t) + o(\Delta t), \quad (4.3)$$

где $p_1(\Delta t)$ — вероятность того, что на отрезке времени Δt завершится одно восстановление.

Для вывода формулы, определяющей $p(x, t)$, на рис. 4.2 изображена поясняющая диаграмма. Событие, вероятность которого равна $p(x, t)$, может произойти двумя способами, а именно:

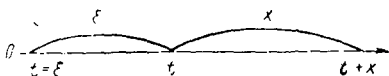


Рис. 4.2. К выводу формулы (4.4)

первый способ — до момента t элемент не подвергается профилактическому обслуживанию с вероятностью $1 - G(t)$ и не отказывает до момента $t+x$ с вероятностью $1 - F(t+x)$; второй способ — в момент $t - \xi$ происходит восстановление элемента с вероятностью $dH(t - \xi)$, и далее в течение времени ξ элемент не подвергается профилактическому обслуживанию с вероятностью $1 - G(\xi)$ и не отказывает за время $\xi + x$ с вероятностью $1 - F(\xi + x)$. Так как ξ меняется в интервале от 0 до t , то каждому значению ξ в данном интервале соответствует определенное значение величин $H(t - \xi)$, $G(\xi)$ и $F(\xi + x)$, поэтому искомая вероятность второго составляющего события находится интегрированием по всем ξ от 0 до t .

Таким образом,

$$p(x, t) = [1 - G(t)] [1 - F(t+x)] + \int_0^t [1 - G(\xi)] [1 - F(\xi+x)] dH(t-\xi). \quad (4.4)$$

Отметим, что в течение фиксированного отрезка времени x (работа в воздухе) отказ элемента возможен, а профилактическое обслуживание исключено, что учтено в выражении (4.4).

Произведем замену переменного во втором слагаемом правой части формулы (4.4): $\xi = t - \omega$. Тогда

$$p(x, t) = [1 - G(t)] [1 - F(t + x)] + \int_0^t [1 - G(t - \omega)] [1 - F(t + x - \omega)] dH(\omega). \quad (4.5)$$

Согласно узловой теореме восстановления (см. гл. I) существует предельное (при $t \rightarrow \infty$) значение вероятности $p(x, t)$, определяемое формулой

$$\lim_{t \rightarrow \infty} p(x, t) = p(x) = \frac{1}{\mu} \int_0^{\infty} [1 - G(\omega)] [1 - F(\omega + x)] d\omega. \quad (4.6)$$

Определим величину μ . Применяя для нахождения искомого условного математического ожидания μ формулу полной вероятности и учитывая, что среднее время работы элемента без отказа и профилактического обслуживания равно

$$\int_0^{\infty} [1 - G(t)] [1 - F(t)] dt, \quad (4.7)$$

имеем

$$\mu = \int_0^{\infty} [1 - G(t)] [1 - F(t)] dt + T_3 \int_0^{\infty} F(x) dG(x) + T_{II} \int_0^{\infty} G(x) dF(x). \quad (4.8)$$

Выражение (4.6) с учетом зависимости (4.8) примет вид

$$p(x) = \frac{\int_0^{\infty} [1 - G(\omega)] [1 - F(\omega + x)] d\omega}{\int_0^{\infty} [1 - G(t)] [1 - F(t)] dt + T_3 \int_0^{\infty} F(x) dG(x) + T_{II} \int_0^{\infty} G(x) dF(x)}. \quad (4.9)$$

Чтобы перейти к определению оптимального периода контроля (профилактического обслуживания), определим функцию $G(t)$ следующим образом:

$$G(t) = \begin{cases} 1, & \text{если } t \geq T_k; \\ 0, & \text{если } t < T_k; \end{cases} \quad (4.10)$$

т. е. предположим, что контроль назначается через постоянное время T_k после включения элемента в работу.

Обоснование того, что наилучшее распределение $G(t)$ является вырожденным, дано в работе [6] и п. 4.5. Тогда

$$p(x) = p(x, T_k) = \frac{\int_0^{T_k} [1 - F(\omega + x)] d\omega}{\int_0^{T_k} [1 - F(t)] dt + T_3 F(T_k) + T_{II} \int_{T_k}^{\infty} dF(x)}. \quad (4.11)$$

При этом мы использовали очевидное равенство

$$\int_{T_k}^{\infty} F(x) dG(x) = F(T_k). \quad (4.12)$$

Вычислим производную выражения (4.11) по T_k , введя обозначения:

$$\int_0^{T_k} [1 - F(\omega)] d\omega = A; \quad F(T_k) = F; \quad f(T_k) = f, \quad (4.13)$$

использовав формулу

$$F(\omega + x) \approx F(\omega) + f(\omega)x \quad (4.14)$$

и считая, что действительное значение x достаточно мало для возможности такой аппроксимации. Тогда

$$\frac{\partial p(x, T_k)}{\partial T_k} = \frac{(1 - F - fx)(A + T_3 F + T_{II} - T_{II} F)}{(A + T_3 F + T_{II} - T_{II} F)^2} - \frac{(A - Fx)(1 - F + T_3 f - T_{II} f)}{(A + T_3 F + T_{II} - T_{II} F)^2}.$$

После проведения элементарных алгебраических преобразований, приравняв производную нулю и пренебрегая членом $T_{II}f \ll 1$, получим

$$\frac{T_{II}}{T_{III} + x} = 1 - \frac{1 - F(T_K)}{[1 - F(T_K)]^2 + f(T_K) \int_0^{T_K} [1 - F(\omega)] d\omega} \quad (4.15)$$

или

$$\frac{T_{II}}{T_{III} + x} = 1 - \frac{1}{1 - F(T_K) + \lambda(T_K) \int_0^{T_K} [1 - F(\omega)] d\omega}. \quad (4.16)$$

Покажем, что при $T_{III} > T_{II}$ и монотонном возрастании $\lambda(t)$ функция $p(x, T_K)$ имеет единственный максимум. Для этого достаточно убедиться в том, что выражение

$$\frac{T_{II}}{T_{III} + x} = 1 + \frac{1}{1 - F(T_K) + \lambda(T_K) \int_0^{T_K} [1 - F(\omega)] d\omega} \quad (4.17)$$

при изменении T_K меняет знак плюс на минус, а это возможно лишь тогда, когда функция

$$\Phi(T_K) = \frac{1}{1 - F(T_K) + \lambda(T_K) \int_0^{T_K} [1 - F(\omega)] d\omega} \quad (4.18)$$

монотонно убывает с возрастанием T_K , т. е. когда ее знаменатель при этом монотонно возрастает. Следовательно, производная от знаменателя выражения (4.18) должна быть положительна:

$$\frac{\partial}{\partial T_K} \left[\frac{1}{\Phi(T_K)} \right] = \lambda'(T_K) \int_0^{T_K} [1 - F(\omega)] d\omega > 0. \quad (4.19)$$

В рассматриваемом случае условие (4.19) всегда выполняется, так как $\lambda(t)$ — монотонно возрастающая функция и производная от нее положительна, а интеграл положителен потому, что интегрируется положительная величина.

При практическом использовании формул (4.15) или (4.16), в которых правая часть зависит только от величины T_K , строят график значений правой части в функции T_K , а по оси ординат откладывают отношение $\frac{T_{II}}{T_{III} + x}$ и

для каждого значения отношения графическим путем определяют оптимальное время контроля (профилактического обслуживания).

После определения величины $T_{конт}$ организация эксплуатации элемента строится так, как показано на рис. 4.3, где светлыми кружками обозначены моменты выполнения плановых замен, а темными — неплановых, обусловленных отказами.

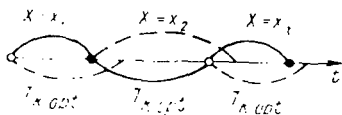


Рис. 4.3. К принципу осуществления замены через неслучайное время

Пример 4.1. Имеется непрерывно работающая система¹ с функцией распределения времени безотказной работы следующего вида:

$$F(t) = 1 - e^{-\frac{qt^2}{2}},$$

где $q = 0,0005 \text{ 1/ч}^2$.

Время выполнения плановых профилактических работ на системе постоянно и равно 15 ч ($T_{п} = 15 \text{ ч}$); среднее время устранения отказа в системе с последующим возвращением ее в исходное состояние равно 20 ч ($T_{з} = 20 \text{ ч}$); оперативное время работы системы 10 ч ($x = 10 \text{ ч}$). Определить из условия обеспечения максимальной готовности системы к работе в произвольный момент времени t и максимальной надежности системы за время от t до $t + 10$ ч интервал выполняемых на ней плановых профилактических работ $T_{к}$, если эксплуатация системы организуется так, как показано на рис. 4.3.

Решение. Убедимся в том, что в нашем случае функция $\lambda(t)$ монотонно возрастает во времени. В самом деле, из выражения (4.1) следует, что $\lambda(t) = qt$.

Далее воспользуемся соотношением (4.16), обозначив правую часть этого соотношения через $\varphi(T_{к})$. Тогда

$$\varphi(T_{к}) = 1 - \frac{1}{e^{-\frac{0,0005T_{к}^2}{2}} + 0,0005T_{к} \int_0^{T_{к}} e^{-\frac{0,0005t^2}{2}} dt}.$$

Расчетные данные функции $\varphi(T_{к})$ для ряда значений ее аргумента сведены в табл. 4.1 и иллюстрируются графиком рис. 4.4.

Таблица 4.1

$T_{к}, \text{ ч}$	8	10	20	30	75	100	150	200
$\varphi(T_{к})$	0	0,03	0,09	0,36	0,57	0,65	0,78	0,83

¹ Здесь система рассматривается как элемент.

По данным табл. 4.1 на рис. 4.4 построен график функции $\varphi(T_K)$, где по оси ординат отложены возможные значения отношения $T_{II}/(T_3+x)$. В рассматриваемом случае

$$\frac{T_{II}}{T_3+x} = \frac{15}{20+10} = 0,5.$$

Следовательно, $T_{K \text{ opt}} \approx 70$ ч.

В некоторых случаях справедлив следующий принцип. Режим сложной системы, при котором максимизируется

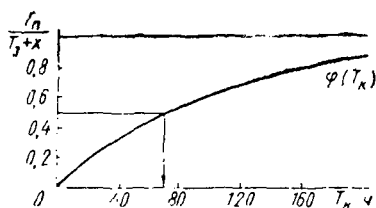


Рис. 4.4. График реализации функции $\varphi(T_K)$

вероятность ее безотказной работы, сводится к максимизации вероятности безотказной работы каждого ее элемента. В этом случае после определения величины $T_{K \text{ opt}}$ для каждого съемного блока или элемента системы (или для регулировок, относящихся ко всей системе) необходимо

проверить значение величины $p_i(x, T_{K \text{ opt}}^{(i)})$, $i=1, \dots, n$, где n — суммарное число съемных блоков (элементов) и общих регулировок системы, или для всей системы, — значение вероятности заставить систему исправной в произвольно выбранный момент времени и проработать безотказно время x , если плановые замены блоков и регулировки системы осуществляются через оптимальные периоды $T_{K \text{ opt}}^{(i)}$, $i=1, \dots, n$.

Обозначим эту вероятность через $P(x, T_{K \text{ opt}}^{(1)}, \dots, T_{K \text{ opt}}^{(n)})$. Очевидно,

$$P(x, T_{K \text{ opt}}^{(1)}, \dots, T_{K \text{ opt}}^{(n)}) = \prod_{i=1}^n r_i(x, T_{K \text{ opt}}^{(i)}).$$

Если системы используются по назначению группой (N систем в группе) и требуется, чтобы в каждый момент времени были готовы к работе не менее qN ($0 < q \leq 1$) систем, то можно ввести, используя биномиальное распределение, следующий критерий надежности и готовности для оценки группового функционирования неза-

висимо работающих систем (при оптимально выбранных плановых сроках замен $T_{\text{к опт}}^{(i)}$, $i = 1, \dots, n$):

$$P_{qN}(x, T_{\text{к опт}}^{(1)}, \dots, T_{\text{к опт}}^{(n)}) = \sum_{j=qN}^N C_N^j \left[\prod_{i=1}^n p_i(x, T_{\text{к опт}}^{(i)}) \right]^j \left[1 - \prod_{i=1}^n p_i(x, T_{\text{к опт}}^{(i)}) \right]^{N-j}. \quad (4.20)$$

Вероятность, определяемая выражением (4.20), при $x=0$ есть доля времени (при длительной эксплуатации группы из N однотипных систем), в течение которого не менее qN систем находятся в исправном состоянии, если осуществляется плановая замена блоков и выполняются регулировки с интервалами $T_{\text{к опт}}^{(i)}$, $i = 1, \dots, n$.

Если выбором величин $T_{\text{к опт}}^{(i)}$, $i = 1, \dots, n$ не удастся обеспечить приемлемое значение критерия $P_{qN}(x, T_{\text{к опт}}^{(1)}, \dots, T_{\text{к опт}}^{(n)})$, то необходимо применять другие меры повышения надежности системы (улучшить ее ремонтпригодность, надежность составляющих систему элементов и схем их соединения).

Рассмотрим теперь, как для САУ реализовать стратегии оптимальных замен блоков и регулировок системы.

Пусть САУ имеет k съёмных блоков и элементов и в ней осуществляется l различных регулировок по обобщенным выходным параметрам ($k+l=n$). В процессе выполнения регулировок возможна замена любого сменного блока (в этом случае считается, что замененные съёмные блоки отказали).

Так как с течением времени (без вмешательства) САУ всегда ухудшает свои свойства, то считаем, что все сменные блоки и регулировки системы во времени характеризуются монотонно возрастающими функциями интенсивностей отказов. Поэтому, применяя изложенные выше методы, можно определить значения $T_{\text{к опт}}^{(i)}$, $i = 1, \dots, k$, $k+1, \dots, n$.

Реализация планов контроля состояния элементов САУ через расчетные сроки $T_{\text{к опт}}^{(i)}$, $i = 1, \dots, n$ (с учетом перепланирования при отказах) зависит от способа эксплуатации системы и организации ее обслуживания, как правило, не допускающих проведения профилактических замен для каждого элемента и выполнения каждой регулировки отдельно. Поэтому всегда оптимальные сроки

замены элементов и выполнения регулировок следует группировать с тем, чтобы при выполнении очередной профилактики системы провести по возможности большее число плановых (оптимальных или близких к ним) замен и регулировок. Это группирование сроков замен и регулировок необходимо корректировать при каждом отказе элементов системы и при каждой разрегулировке.

Группирование сроков замен элементов и регулировок на практике всегда возможно вследствие того, что функции $p_i(x, T_k^{(i)})$, ($i=1, \dots, k, \dots, n$) имеют, как правило, пологий участок в окрестности максимума, поэтому плановые замены элементов или регулировок проводят в интервалах $T_{k \text{ опт}}^{(i)} \pm \Delta T_i$, причем величина ΔT_i определяется при решении конкретных задач.

После осуществления очередной группировки сроков устанавливается единый срок профилактики системы, для которого однозначно определен объем профилактики, обуславливающий проведение запланированных оптимальных замен и регулировок.

Для удобства планирования предупредительных замен элементов и проведения профилактических регулировок рекомендуется заполнять таблицу (см. табл. 4.2).

Для заполнения таблицы необходимо провести подготовительную и вычислительную работу: собрать статистический материал по отказам блоков (элементов) и разрегулировкам параметров системы и построить функции $\lambda_i(t)$, $i=1, \dots, k, \dots, n$, проверив статистические гипотезы о возрастании этих функций во времени. Статистический материал собирается в процессе проведения стендовых испытаний блоков (элементов) системы и при испытаниях всей системы, он также может быть частично получен и из опыта эксплуатации аналогичных элементов, блоков, работающих в другой аппаратуре в условиях, близких к предполагаемым условиям работы элементов и блоков проектируемой САУ.

После построения функций $\lambda_i(t)$, заполнения колонки 5 табл. 4.2 и построения функций $\varphi(T_{ик})$ с использованием правой части выражения (4.16), можно заполнить основную колонку таблицы — колонку 2.

Желанием конструктора, планирующего замены элементов на период эксплуатации проектируемой системы, является знание коэффициентов q_i ($i=1, \dots, k, \dots, n$), указывающих, какую часть от среднего времени безотказной работы блока (элемента) или среднего времени между

профилактическими регулировками параметра составляет оптимальный период предупредительной замены блока (элемента) или оптимальный период упреждающей профилактической регулировки параметра соответственно.

Таблица 4.2

№ заменяемых блоков и регулируемых параметров	$T_{i\text{к opt}}$	Среднее время безотказной работы блока или среднее время между разрегулировками параметра $T_{i\text{o}}$	$q_i = \frac{T_{i\text{к opt}}}{T_{i\text{o}}}$	$\frac{T_{i\text{п}}}{T_{i\text{з}+x}}$
1	2	3	4	5
1	$T_{1\text{к opt}}$	$T_{1\text{o}}$	q_1	$\frac{T_{1\text{п}}}{T_{1\text{з}+x}}$
2	$T_{2\text{к opt}}$	$T_{2\text{o}}$	q_2	$\frac{T_{2\text{п}}}{T_{2\text{з}+x}}$
·	·	·	·	·
·	·	·	·	·
·	·	·	·	·
k	$T_{k\text{к opt}}$	$T_{k\text{o}}$	q_k	$\frac{T_{k\text{п}}}{T_{k\text{з}+x}}$
l {	$k+1$	$T_{k+1\text{к opt}}$	q_{k+1}	$\frac{T_{k+1\text{п}}}{T_{k+1\text{з}+x}}$
	·	·	·	·
	·	·	·	·
	·	·	·	·
n	$T_{n\text{к opt}}$	$T_{n\text{o}}$	q_n	$\frac{T_{n\text{п}}}{T_{n\text{з}+x}}$

Коэффициенты q_i могут быть получены и сведены в стандартные таблицы при проведении массовых расчетов по описанному алгоритму (для каждого фиксированного значения $i=1, \dots, k, \dots, n$), т. е. для всех одинаковых блоков, элементов и систем, работающих в составе различных систем и комплексов в условиях, близких к предполагаемым условиям работы проектируемой системы. В этом случае знания величины q_i и $T_{i\text{o}}$ ($i=1, \dots, k, \dots, n$) будет достаточно для определения значений $T_{i\text{к opt}}$, однако, следует иметь в виду, что по мере сбора статистических данных в процессе будущей эксплуатации систем значения $T_{i\text{к opt}}$ должны уточняться с использованием изложенного выше алгоритма.

Пример 4.2¹. Рассмотрим СНС стабилизации канала крена (см. гл. II), при несколько измененных исходных данных. Предположим, что эта система должна эксплуатироваться в течение 2000 ч по наработке. Определим для СНС оптимальные интервалы замен съемных элементов и необходимый объем ЗИП. Предположим, что СНС сконструирована так, что состоит из 10 съемных элементов с интенсивностями отказов и значениями $T_{п}$, $T_{з}$, приведенными в табл. 4.3; $x=5$ ч. В этой же таблице приведены рассчитанные значения $T_{i \text{ опт}}$, $i=1, \dots, 10$.

Среднее число отказов каждого элемента автомата в течение 2000 ч (ЗИП без предупредительной замены) при отсутствии упреждающих замен и суммарное среднее число отказов и упреждающих оптимальных профилактических замен для случая замены элементов (ЗИП с оптимальной заменой) представлены в табл. 4.4.

Таблица 4.3

Элемент системы	$\lambda(t)$, $1/\text{ч}$	$T_{п}$, ч	$T_{з}$, ч	$\frac{T_{п}}{T_{з}+x}$	$T_{к \text{ опт}}$, ч
1	$0,006 + 0,002t$	3	6	0,272	20
2	$0,004 + 0,015t$	3	6	0,272	22
3	$0,002 + 0,015t$	4	5	0,4	31
4	$0,008 + 0,01t$	4	14	0,21	23
5	$0,0015 + 0,0005t$	4	5	0,4	55
6	$0,001 + 0,0005t$	4	12	0,235	36
7	$0,004 + 0,0004t$	3	5	0,3	51
8	$0,002 + 0,0003t$	4	5	0,4	72
9	$0,0006 + 0,0002t$	4	12	0,235	56
10	$0,0008 + 0,00008t$	12	13	0,67	203

Таблица 4.4

Элемент системы	ЗИП, шт.		% упрежденных отказов	Элемент системы	ЗИП, шт.		% упрежденных отказов
	без предупредительной замены	с оптимальной заменой			без предупредительной замены	с оптимальной заменой	
1	62	98	33	6	38	52	46
2	50	88	58	7	40	51	33
3	52	68	44	8	28	30	31
4	48	78	60	9	22	35	47
5	34	49	30	10	14	15	8

¹ Расчеты с использованием моделирования выполнены инж. В. И. Ивановым.

В обоих случаях среднее число замен определялось как

$$\frac{2000}{\mu_{qi}} \cdot \alpha, \quad i = 1, \dots, 10; \quad q = 1, 2, \quad \text{где } \alpha \ll \frac{2000}{\mu_{qi}}; \quad q = 1$$

соответствует отсутствию замен, а $q=2$ — их наличию; вероятность того, что выбранного ЗИП хватит для проведения замен, принималась равной единице.

Как видно из табл. 4.4, процент упрежденных внезапных отказов съёмных элементов доходит до 60 при осуществлении оптимальной политики упреждающих замен; число элементов в ЗИП при этом возрастает примерно в 1,5—2 раза.

Из изложенного выше можно сформулировать некоторые важные рекомендации по конструированию и обслуживанию САУ.

1. При проектировании САУ необходимо предусматривать (по возможности) конструкции со съёмными блоками. Число съёмных блоков зависит от их надёжности и от того, какое рассогласование в работу САУ вносит замена блока.

2. Необходимо предусмотреть увеличение комплектов ЗИП примерно в 1,5—2 раза по сравнению с нормами, при которых не учитываются упредительные замены.

3. Реализация метода оптимальных упреждающих профилактических замен съёмных элементов позволяет уменьшить число внезапных отказов системы в процессе эксплуатации на 30—60%.

4.2. ПЛАНИРОВАНИЕ ЗАМЕНЫ ЭЛЕМЕНТОВ СИСТЕМЫ ПРИ ОТСУТСТВИИ МГНОВЕННОЙ ИНДИКАЦИИ ОТКАЗОВ

В процессе эксплуатации САУ при конечном времени контроля систем может возникнуть задача выбора периода контроля состояния отдельных элементов в условиях, когда они не проверяются в промежутках времени между моментами контроля. Если предположить, что САУ не может быть использована по назначению при контроле ее технического состояния, то частый контроль снижает готовность системы к работе. Однако и достаточно редкий контроль также приводит к снижению ее готовности, так как в этом случае большое время САУ не может быть использована по назначению по причине отказа элементов (отказ системы выявляется и устраняется только при проведении контроля). Очевидно, и в рассматриваемой ситуации должен существовать оптимальный период контроля системы.

Описанная физическая задача типична для систем, недоступных частым проверкам или непрерывному контролю (отсутствие встроенных индикаторов, недоступность проверок со стороны оператора и т. п.).

Перейдем к математической постановке задачи. Предположим, что время контроля исправного элемента равно V , неисправного — U , время от момента окончания контроля до отказа — X , а время между последовательными моментами контроля (от момента окончания предыдущего контроля до начала последующего контроля) — Y .

Будем вначале считать, что все эти времена являются независимыми случайными величинами с соответствующими функциями распределения: $F_V(t)$, $F_U(t)$, $F(t)$, $G(t)$.

Обозначим через Z следующую величину:

$$Z = \begin{cases} Y + V, & \text{если } X > Y; \\ Y + U, & \text{если } X < Y. \end{cases} \quad (4.21)$$

Основное допущение вновь будет состоять в том, что величины $\{Z_k\}$ образуют процесс восстановления с $\mu = M[Z]$ и функцией восстановления $H(t) = M[N_t]$. Тогда оказывается верным выражение (4.6), где, по-прежнему, $\mu = M[Z]$.

Определим для данного случая величину μ . Очевидно,

$$\mu = M \begin{cases} Y + V, & \text{если } X > Y; \\ Y + U, & \text{если } X < Y. \end{cases} \quad (4.22)$$

Будем считать, что:

$$F_V(t) = \begin{cases} 1, & \text{если } t \geq T_V; \\ 0, & \text{если } t < T_V; \end{cases} \quad (4.23)$$

$$F_U(t) = \begin{cases} 1, & \text{если } t \geq T_U; \\ 0, & \text{если } t < T_U; \end{cases} \quad (4.24)$$

$$G(t) = \begin{cases} 1, & \text{если } t \geq T_K; \\ 0, & \text{если } t < T_K. \end{cases} \quad (4.25)$$

Применяя формулу полной вероятности и учитывая выражения (4.22) — (4.25), а также то, что среднее время между заменами элемента в модели

$$T_K = \int_0^{\infty} [1 - G(t)] dt,$$

$$\mu = T_K + T_U \int_{T_K}^{\infty} F(t) dG(t) + T_V \int_{T_K}^{\infty} dF(t) = \\ = T_K + T_U F(T_K) + T_V [1 - F(T_K)]. \quad (4.26)$$

С учетом изложенного выше выражение (4.6) переписывается следующим образом:

$$p(x, T_K) = \frac{\int_0^{T_K} [1 - F(\omega + x)] d\omega}{T_K + T_U F(T_K) + T_V [1 - F(T_K)]}. \quad (4.27)$$

Возьмем производную от логарифма выражения (4.27) по T_K и приравняем ее нулю. После чего получим

$$\frac{1 - F(x + T_{K \text{ opt}})}{\int_0^{T_{K \text{ opt}}} [1 - F(x + \omega)] d\omega} = \frac{(T_U - T_V) f(T_{K \text{ opt}}) + 1}{T_{K \text{ opt}} + (T_U - T_V) F(T_{K \text{ opt}}) + T_V},$$

ИЛИ

$$T_{K \text{ opt}} [1 - F(x + T_{K \text{ opt}})] + (T_U - T_V) F(T_{K \text{ opt}}) \times \\ \times [1 - F(x + T_{K \text{ opt}})] + T_V - T_V F(x + T_{K \text{ opt}}) = \\ = \int_0^{T_{K \text{ opt}}} [1 - F(x + \omega)] d\omega [T_U - T_V] f(T_{K \text{ opt}}) + 1], \quad (4.28)$$

где $f(t) = F'(t)$.

Существование решения уравнения (4.28) очевидно из качественного анализа поведения его левой и правой частей в точках $T_K = 0$ и $T_K = \infty$.

На практике следует в каждом случае отыскивать решение уравнения (4.28) графически, вводя приемлемые допущения, основанные на малости времени x , ограниченности величины T_K и т. д.

Пример 4.3. Имеем элемент, который в любой заранее неизвестный момент времени t без специально организованной проверки может быть использован по назначению.

Дано: $F(t) = 1 - e^{-\lambda t}$.

Необходимо выбрать период проверки элемента так, чтобы максимизировать вероятность, определяемую выражением (4.27). Будем считать для простоты, что $x = 0$.

Решение. Из уравнения (4.27) находим:

$$\frac{e^{-\lambda T_{\text{к опт}}}}{\frac{1}{\lambda} - \frac{1}{\lambda} e^{-\lambda T_{\text{к опт}}}} = \frac{(T_U - T_V) \lambda e^{-\lambda T_{\text{к опт}}} + 1}{T_{\text{к опт}} + (T_U - T_V)(1 - e^{-\lambda T_{\text{к опт}}}) + T_V}.$$

Из анализа поведения правой и левой частей полученного уравнения видно, что его решение является единственным.

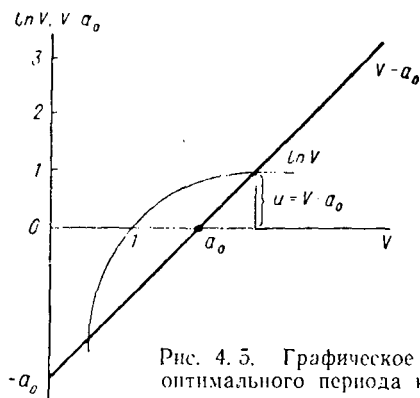


Рис. 4.5. Графическое определение оптимального периода контроля САУ

После элементарных преобразований получаем

$$e^{-\lambda T_{\text{к опт}}} \left(T_{\text{к опт}} + T_V + \frac{1}{\lambda} \right) - \frac{1}{\lambda} = 0,$$

или

$$e^{-u}(u + a_0) - 1 = 0, \quad (4.29)$$

где

$$u = \lambda T_{\text{к опт}}; \quad a_0 = \lambda T_V + 1; \quad a_0 > 1.$$

Пусть далее $V = u + a_0$ ($V > 1$). Тогда $u = V - a_0$ и, логарифмируя выражение (4.29), получим

$$-(V - a_0) + \ln V = 0,$$

или

$$\ln V = V - a_0 \quad (V > 1). \quad (4.30)$$

Так как $\lambda T_{\text{к опт}} = u = V - a_0$, то искомый оптимальный период замены равен

$$T_{\text{к опт}} = \frac{V - a_0}{\lambda}, \quad (4.31)$$

где разность $V - a_0$ непосредственно считывается с графика (рис. 4.5).

4.3. ПЛАНИРОВАНИЕ ЗАМЕНЫ ЗАРЕЗЕРВИРОВАННЫХ ЭЛЕМЕНТОВ СИСТЕМЫ В СЛУЧАЕ ИНДИКАЦИИ ОТКАЗОВ ПРИ ПРОВЕРКАХ

Рассмотрим зарезервированный элемент (ЗЭ), состоящий из идентичных элементов, которые подразделяются на m основных и $n-m$ резервных элементов (рис. 4.6). Зарезервированный элемент работает непрерывно за исключением времени, отводимого на проведение запланированных профилактических мероприятий.

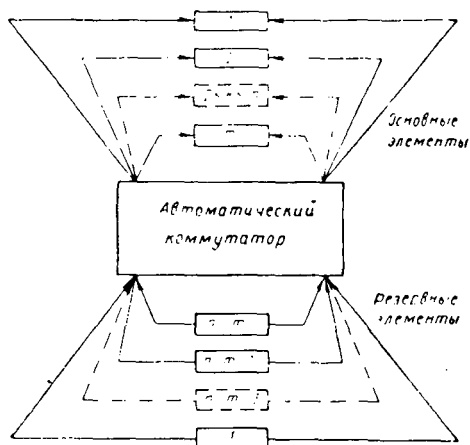


Рис. 4.6. Структура зарезервированного элемента

Возможными состояниями ЗЭ к моменту проведения профилактических мероприятий являются либо исправная работа всех n элементов (в этом случае на проверку технического состояния элементов, а при необходимости и на регулировку их параметров в среднем затрачивается время T_0), либо исправная работа $n-1$ элемента и отказ одного основного (в этом случае время профилактических мероприятий в среднем равно T_1), . . . , либо отказ всех n элементов ЗЭ (при этом время профилактического обслуживания в среднем равно T_n , причем $T_n > T_{n-1} > \dots > T_1 > T_0$). Каждый раз после проведения профилактических мероприятий ЗЭ возвращается в первоначальное состояние.

В процессе работы ЗЭ восстановлений или замен отказавших элементов не производится, а осуществляется лишь автоматическое подключение резервных элементов вместо отказавших основных. Зарезервированный элемент считается отказавшим при выходе из строя одного из основных элементов и при отсутствии исправных резервных элементов.

Задача заключается в определении оптимального периода T_k профилактического обслуживания ЗЭ, при назначении которого обеспечивается максимальная вероятность застать ЗЭ в исправном состоянии в произвольный момент времени t (максимальная готовность) и максимальная вероятность безотказной работы ЗЭ за время от t до $t+x$ (максимальная надежность).

Обозначим, как и ранее, вероятность того, что ЗЭ работает к моменту t и продолжает работать в течение времени от t до $t+x$, через $p(x, t)$, а функцию распределения времени безотказной работы ЗЭ (случайной величины X), через $F(t) = P\{X < t\}$.

Вначале предположим, что время между двумя последовательными моментами профилактического обслуживания ЗЭ является случайной величиной Y с функцией распределения $G(t) = P\{Y < t\}$.

Так как после каждого профилактического обслуживания ЗЭ возвращается в первоначальное состояние, то времена между двумя моментами окончания профилактических обслуживаний Z_k образуют вновь процесс восстановления.

Величина Z может принимать следующие значения:

$$Z = \begin{cases} Y + T_0, & \text{если } k=0; \\ Y + T_1, & \text{если } k=1; \\ Y + T_2, & \text{если } k=2; \\ \dots & \dots \\ Y + T_n, & \text{если } k=n, \end{cases}$$

где $k=1, 2, \dots, n$ — количество отказавших элементов к началу профилактического обслуживания.

В рассматриваемом случае также справедлива формула (4.6), в которой

$$\mu = M[z] = M[Y] + p_0(Y)T_0 + p_1(Y)T_1 + \dots + p_n(Y)T_n,$$

причем $\sum_{i=0}^n p_i(Y) = 1$ и $p_i(Y)$, $i=0, 1, 2, \dots, n$ обозначает вероятность к началу профилактического обслуживания иметь i отказавших элементов.

Выберем $G(t)$ по отношению к T_K так, чтобы выполнялось условие (4.10). Тогда на основании выражения (4.6)

$$p(x, T_K) = \frac{\int_0^{T_K} [1 - F(\omega + x)] d\omega}{T_K + p_0(T_K)T_0 + p_1(T_K)T_1 + \dots + p_n(T_K)T_n}. \quad (4.32)$$

Формула (4.32) дает возможность выбирать период T_K профилактического обслуживания ЗЭ, исходя из заданной величины $p(x, T_K)$. Однако, наибольший интерес представляет отыскание оптимальной величины T_K , обеспечивающей максимум функции $p(x, T_K)$ и определяемой из условий

$$\frac{\partial p(x, T_K)}{\partial T_K} = 0 \quad (4.33)$$

при существовании (для каждого конкретного случая) единственности максимума $p(x, T_K)$.

Для определенности рассмотрим ЗЭ (а для получения окончательных результатов перейдем впоследствии к еще более частным случаям), состоящий из одного основного элемента ($m=1$) и $n-1$ резервных элементов, которые находятся в состоянии ненагруженного резерва. Тогда время безотказной работы ЗЭ

$$X = X_1 + X_2 + \dots + X_n, \quad (4.34)$$

где X_1 — время безотказной работы основного элемента;
 X_2 — время безотказной работы первого резервного элемента;
 X_n — время безотказной работы $(n-1)$ -го резервного элемента.

Следовательно,

$$Z = \left\{ \begin{array}{l} Y + T_0, \text{ если } X_1 > Y; \\ Y + T_1, \text{ если } X_1 < Y, X_1 + X_2 > Y; \\ Y + T_2, \text{ если } X_1 + X_2 < Y, X_1 + X_2 + X_3 > Y; \\ \dots \\ Y + T_n, \text{ если } X < Y. \end{array} \right\} \quad (4.35)$$

Очевидно, для случая ненагруженного резерва

$$F(t) = 1 - \left[1 + \lambda t + \frac{(\lambda t)^2}{2!} + \dots + \frac{(\lambda t)^{n-1}}{(n-1)!} \right] e^{-\lambda t} \quad (4.36)$$

при условии, что интенсивность отказов каждого элемента остается постоянной и равной λ , а характер потока отказов элемента простейший. В этом случае выражение (4.32) принимает вид

$$P(x, T_k) = \frac{\int_0^{T_k} \left\{ 1 + \lambda(t+x) + \frac{[\lambda(t+x)]^2}{2!} + \dots + \frac{[\lambda(t+x)]^{n-1}}{(n-1)!} \right\} e^{-\lambda(t+x)} dt}{T_k + T_0 e^{-\lambda T_k} + T_1 \lambda T_k e^{-\lambda T_k} + \dots + T_{n-1} \frac{(\lambda T_k)^{n-1}}{(n-1)!} e^{-\lambda T_k} + T_n \left\{ 1 - \left[1 + \lambda T_k + \dots + \frac{(\lambda T_k)^{n-1}}{(n-1)!} \right] e^{-\lambda T_k} \right\}} \quad (4.37)$$

Рассмотрим для примера случай: $m=1$, $n=2$, т. е. когда в состав системы входят один основной и один резервный элемент и для каждого элемента поток отказов простейший с параметром λ . Тогда

$$p(x, T_k) = \frac{\int_0^{T_k} \left\{ 1 + \lambda(t+x) + \frac{[\lambda(t+x)]^2}{2} \right\} e^{-\lambda(t+x)} dt}{T_k + T_0 e^{-\lambda T_k} + T_1 \lambda T_k e^{-\lambda T_k} + T_2 \{ 1 - [1 + \lambda T_k] e^{-\lambda T_k} \}} \quad (4.38)$$

Убеждаясь в существовании единственного максимума функции $p(x, T_k)$ и используя условие (4.33), получаем следующее уравнение:

$$e^{-\lambda T_k} [2T_k (1 + \lambda T_1) + T_2 (1 - \lambda T_k) - 2(T_1 - T_0) + \lambda \left(T_k^2 + \frac{2}{\lambda^2} \right)] - e^{-2\lambda T_k} (T_0 - 2T_1 + T_2) - \frac{2}{\lambda} = 0. \quad (4.39)$$

Уравнение (4.39) легко решается графически (рис. 4.7). При этом удобно все члены, зависящие от T_k , обо-

значить через функцию $\Phi(T_K)$. В табл. 4.5 приведены результаты вычислений значений $T_{K \text{ опт}}$ для случаев $m=1$; $n=2, 3, 4$ при различных значениях λ . Вычисления производились с использованием выражения (4.37) и условия (4.33).

Ясно, что при $n > 2$ значительно усложняются выкладки при определении оптимального значения величины T_K , однако, принципиальных затруднений не встречается.

В более сложных случаях (нагруженный, облегченный резерв, а также при $\lambda \neq \text{const}$) вероятности p_i могут быть определены с помощью биномиального распределения (нагруженный резерв) или при применении так называемой «схемы гибели и размножения» к исследованиям резервных систем, которое было предложено и обосновано А. Д. Соловьевым.

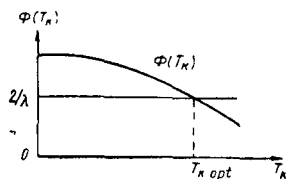


Рис. 4.7. Графическое определение оптимального периода контроля САУ с резервом

Таблица 4.5

$\lambda \cdot 10^4$	$T_0, \text{ ч}$	$T_1, \text{ ч}$	$T_2, \text{ ч}$	$T_3, \text{ ч}$	$T_4, \text{ ч}$	$T_{K \text{ опт}}, \text{ ч}$		
						$n=2$ $m-1$	$n=3$ $m-1$	$n=4$ $m-1$
0,01	0,2	0,4	0,6	0,8	1	8,7	12,7	17,3
0,02	0,3	0,6	0,9	1,2	1,5	14,1	23,8	32,9
0,01	0,5	1	1,5	2	2,5	26	45,9	63,6
0,005	0,7	1,5	2,4	3,2	4	45,9	83,6	117,5
0,0025	1	2	3	4	5	81,8	155,9	232,2
0,00125	2	3	4	5	6	163,7	316,5	488,7
0,00066	4	5,5	7	8,5	10	316	610,2	945
0,00033	6	7,5	9	10,5	12	572	1138	1823,4
0,00016	8	10	12	14	16	1014,9	2097,3	3169,3
0,000083	10	12	14	16	18	1685,8	3617,1	6174,7

4.4. ОПТИМАЛЬНОЕ ОБСЛУЖИВАНИЕ СИСТЕМЫ ПО НАРАБОТКЕ И СОСТОЯНИЮ СОСТАВЛЯЮЩИХ ЭЛЕМЕНТОВ

Опишем сначала кратко модель, подробно изложенную в работе [6]. Рассмотрим, пренебрегая простоями системы, ее непрерывную работу.

Обозначим через $v(t)$ параметр, характеризующий состояние системы в процессе работы:

$$v(t) = \begin{cases} 0, & \text{если система исправна в момент } t; \\ 1, & \text{в противном случае.} \end{cases}$$

Пусть $\theta(t)$ — наработка системы к произвольному моменту t , тогда состояние системы в момент t может быть описано вектором $\{v(t), \theta(t)\}$. Считаем, что в области A возможных состояний систем известна подобласть A_0 : при $\{v(t), \theta(t)\} \in A_0$ система считается отказавшей. Пусть далее в момент $t_0=0$ начинается эксплуатация новой системы.

Предполагается, что в моменты $t_i=i\Delta t$ ($i=1, 2, \dots$) наблюдается состояние системы и определяется вектор $\{v(t_i), \theta(t_i)\}$. Если в момент t_i ($i=1, 2, \dots$) обнаруживается, что система ($\{v(t_i), \theta(t_i)\} \in A_0$) отказала, то она подвергается аварийно-профилактическому ремонту. В противном случае либо принимается решение о проведении предупредительной профилактики системы, либо о продолжении ее эксплуатации до следующего момента контроля $t_{i+1}=t_i+\Delta t$ и т. д. Считаем, что после профилактики или ремонтов система является работоспособной и ее наработка равна нулю.

Если принять, что аварийно-профилактический ремонт системы занимает постоянное время $\gamma_{a,n}$ с коэффициентом пропорциональности по стоимости $c_{a,n}$, то стоимость его проведения

$$\tilde{c}_{a,n} = c_{a,n} \gamma_{a,n}.$$

Аналогично для стоимости $\tilde{c}_{п,n}$ проведения предупредительной профилактики

$$\tilde{c}_{п,n} = c_{п,n} \gamma_{п,n}.$$

Считаем, что $\tilde{c}_{a,n} > \tilde{c}_{п,n}$. Стоимость проверки работоспособной системы (контроля ее состояния) полагаем равной нулю. Моменты окончания профилактики и ремонтов являются моментами регенерации (обновления) процесса, описывающего состояние системы. Интервал времени между соседними моментами регенерации назовем периодом регенерации.

Искомая стратегия обслуживания в период регенерации полностью определяется разбиением области A возможных состояний системы $\{v(t_k), t_k\}$ на три непересекающиеся подобласти A_1, A_2, A_3 ($A_1 \cup A_2 \cup A_3 = A$), при попадании в которые вектора $\{v(t_k), t_k\}$ принимается

решение о продолжении наблюдений, о проведении предупредительного профилактического мероприятия или аварийно-профилактического ремонта соответственно ($A_3=A_0$). Наша задача заключается в оптимальном разбиении областей A_1 и A_2 .

Введем критерий оптимальности разбиения области A . Пусть произведено некоторое разбиение A_1, A_2, A_3 и на рассматриваемом периоде регенерации время пребывания системы в области A_1 равно t . Тогда удельные затраты за период регенерации равны:

$$c(A_1, A_2, A_3) = c_t = \begin{cases} \frac{\tilde{c}_{\text{п.п}}}{t}, & \text{если в момент } t \text{ осуществ-} \\ \text{лен переход в область } A_2; \\ \frac{\tilde{c}_{\text{а.п}}}{t}, & \text{если в момент } t \text{ осуществ-} \\ \text{лен переход в область } A_3 = A_0. \end{cases} \quad (4.40)$$

Время t пребывания системы в области A_1 зависит от разбиения A_1, A_2, A_3 . За критерий, характеризующий качество разбиения A_1, A_2, A_3 , а значит, и качество стратегии обслуживания в работе [6] были приняты средние удельные затраты, приходящиеся на единицу времени безотказной работы системы за период регенерации:

$$Mc_t = \tilde{c}(A_1, A_2, A_3).$$

Задача теперь свелась к определению таких областей $A_1^*, A_2^*, A_3^* = A_0 (A_1^* \cup A_2^* \cup A_3^* = A)$, для которых достигается минимум средних удельных затрат:

$$\tilde{c}(A_1^*, A_2^*, A_3^*) = \min_{\substack{A_1, A_2, A_3 \\ A_1 \cup A_2 \cup A_3 = A}} \tilde{c}(A_1, A_2, A_3). \quad (4.41)$$

Строго говоря, критерием (4.41) можно пользоваться при многократном применении формулируемого ниже оптимального правила разбиения A_1, A_2, A_3 в процессе длительной эксплуатации системы только в случае, когда возможные изменения величины t невелики. В остальных случаях управление не будет близким к оптимальному, хотя и может быть использовано на практике. Строгая оптимальность правила соблюдается при эксплуатации систем до первого попадания в области A_2 или A_3 .

Покажем, когда можно предложенным правилом оптимального разбиения пользоваться на практике для систем с длительным периодом эксплуатации. Пусть T — период эксплуатации системы, ξ — случайная длина периода регенерации ($M\xi = \tau$), φ — математическое

ожидание затрат η на периоде регенерации ($M\eta = \varphi$), D — множество стратегий разбиения d . Тогда суммарные затраты за время T

$$f(T) = \sum_{i=1}^{N(T)} \eta_i + \eta_{N(T)+1}^\theta, \quad 0 \leq \theta \leq 1,$$

где $N(T)$ — число периодов регенерации до момента T .

Если $\frac{f(T)}{T} \underset{T \rightarrow \infty}{>} \gamma$, то нам нужно получить $\min_{d \in D} \gamma = \gamma_0$.

Ясно, что $\frac{N(T)}{T} \rightarrow \frac{1}{\tau}$, а $\frac{f(T)}{T} = \frac{f(T)}{N(T)} \frac{N(T)}{T} \rightarrow \frac{M\eta}{M\xi}$

$$\text{и} \quad \left\{ \min_{d \in D} \gamma \right\} \rightarrow \left\{ \min \frac{M\eta}{M\xi} \right\}.$$

Нас интересует, когда можно считать, что $M \frac{\eta}{\xi} \approx \frac{M\eta}{M\xi}$. С помощью теории выпуклых фигур (применяя и рандомизированные правила разбиения) можно показать, что при некотором ϱ

$$\left\{ \min \frac{M\eta}{M\xi} \right\} \rightarrow \left\{ \min M[\eta(d) - \varrho\xi(d)] \right\}.$$

Если $P\{a_0 \leq \xi \leq a_1\} = 1$ и $P\{b_0 \leq \eta \leq b_1\} = 1$, где $b_0 \geq \varphi - \Delta\eta$; $b_1 \leq \varphi + \Delta\eta$; $a_0 \geq \tau - \Delta\xi$; $a_1 \leq \tau + \Delta\xi$, то для параметра ϱ , обеспечивающего искомый \min (так как $\max \left\{ \frac{\eta}{\xi} \right\} = \frac{b_1}{a_0}$, а $\min \left\{ \frac{\eta}{\xi} \right\} = \frac{b_0}{a_1}$), можно написать, что

$$\frac{M\eta}{M\xi} \left(1 - \frac{\Delta\eta}{M\eta} - \frac{\Delta\xi}{M\xi} \right) \leq M \frac{\eta}{\xi} \leq \frac{M\eta}{M\xi} \left(1 + \frac{\Delta\eta}{M\eta} + \frac{\Delta\xi}{M\xi} \right).$$

Последнее выражение позволяет обосновать возможность перехода от критерия $M \frac{\eta}{\xi}$ к критерию $\frac{M\eta}{M\xi}$, т. е. использования полученных выше результатов для систем длительной эксплуатации.

Для определения оптимального разбиения A_1^* , A_2^* , A_3^* , обеспечивающего минимум средним удельным затратам $\tilde{c}(A_1, A_2, A_3)$, воспользуемся леммой Дуба (см. п. 1.6). В работе [6] эта лемма применительно к рассматриваемому случаю сформулирована следующим образом.

Пусть $\{F_s\}$, $s=1, 2, \dots$ — последовательность σ -полей в некотором пространстве W элементарных событий, $F_s \subset F_{s+1}$, а $\{y_s\}$ — последовательность случайных величин, измеримых относительно F_s , и для которых существует $M[y_t]$. Пусть также задан некоторый класс K правил определения моментов k остановки наблюдений таких, что $M[k] < \infty$.

Если существует правило, определяющее момент остановки k^* , для которого:

$$1) M[k^*] < \infty;$$

$$2) M\{y_s | F_{s-1}\} \begin{cases} \leq y_{s-1}, & s \leq k^*; \\ \geq y_{s-1}, & s > k^* \end{cases}$$

для почти всех элементарных событий $\omega \in W$, где s — моменты остановок наблюдений, определяемые другими правилами из класса K , и если существует константа $C < \infty$ такая, что для всех s

$$3) M\{|y_{s+1} - y_s| | F_s\} \leq C,$$

то данное правило является оптимальным, т. е.

$$M[y_{k^*}] = \min M[y_s],$$

где минимум берется по всем возможным правилам остановки из класса K .

В работе [6] было показано, что условия «1)» и «3)» леммы на практике всегда выполняются. Относительно выполнения условия «2)» необходимо сказать следующее. Для реальной технической системы, не подвергающейся обслуживанию, естественным является предложение о том, что с увеличением наработки надежность системы уменьшается. В этом случае вероятность

$$P\{[v(t_k), t_k] \in A_3 | F_{t_{k-1}}\}$$

является монотонно возрастающей функцией времени $t_k = k\Delta t$, что необходимо и достаточно для выполнения условия «2)», которое в рассматриваемом случае записывается так:

$$M\{c_{t_k} | F_{t_{k-1}}\} \begin{cases} \leq c_{t_{k-1}}, & \text{если } P\{[v(t_k), t_k] \in A_3 \\ & \text{при } [v(t_i), t_i, i = \overline{0, k-1}]\} \leq \\ & \leq \frac{\tilde{c}_{н.н}}{(\tilde{c}_{а.н} - \tilde{c}_{н.н})(k-1)}; \\ \geq c_{t_{k-1}}, & \text{если } P\{[v(t_k), t_k] \in A_3 \\ & \text{при } [v(t_i), t_i, i = \overline{0, k-1}]\} \geq \\ & \geq \frac{\tilde{c}_{н.н}}{(\tilde{c}_{а.н} - \tilde{c}_{н.н})(k-1)}. \end{cases} \quad (4.42)$$

В данном случае последовательность $\{F_{t_s}\}$ есть последовательность наблюдаемых состояний системы

$$F_{t_s} = \{v(t_i), t_i, i = \overline{0, s}\},$$

а момент оптимальной остановки [см. выражение (4.42)] определяется с помощью неравенства

$$P \{ [v(t_k), t_k] \in A_3 | [v(t_i), t_i, i = \overline{0, k-1}] \} \leq \leq \frac{\tilde{c}_{п.п}}{(\tilde{c}_{а.п} - \tilde{c}_{п.п})(k-1)}. \quad (4.43)$$

Правая часть выражения (4.43) монотонно убывает, а левая возрастает при возрастании k . Поэтому, если в качестве k^* взять такое k , для которого последний раз выполняется неравенство (4.43), то условие 2 будет выполняться.

Таким образом, все условия леммы Дуба применительно к данному случаю оказываются выполненными, и оптимальное правило остановки (нахождения оптимальной границы между областями A_1 и A_2) определяется с помощью выражения (4.43), которое в случае оптимального обслуживания системы по наработке принимает вид:

$$1 - \frac{e^{-\int_0^{t_{k-1} + \Delta t} \Lambda(t) dt}}{e^{-\int_0^{t_{k-1}} \Lambda(t) dt}} = 1 - e^{-\int_{t_{k-1}}^{t_k} \Lambda(t) dt} \leq \leq \frac{\tilde{c}_{п.п}}{(\tilde{c}_{а.п} - \tilde{c}_{п.п})(k-1)}, \quad (4.44)$$

где $\Lambda(t)$ — функция интенсивности отказов системы, определяемая выражением (4.1).

Правило построения области A_1 таково: всякая точка $\{v(t_{k-1}, t_k)\}$, для которой выполняется (4.44), принадлежит области A_1 .

Обслуживание системы по состоянию составляющих ее элементов

Параллельно - последовательное соединение элементов

На практике во многих случаях соединение элементов системы по надежности можно представить как ком-

бинированное соединение параллельных и последовательных элементов.

Рассмотрим параллельно-последовательное соединение элементов системы по надежности, представленное на рис. 4.8. Пусть число последовательно соединенных разных по надежности элементов системы равно N , а число работающих параллельно и также разных по надежности элементов — n . Отказ системы (предполагаем, что она работает непрерывно) наступает при отказе

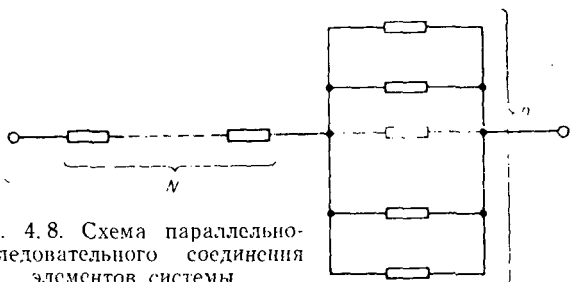


Рис. 4.8. Схема параллельно-последовательного соединения элементов системы

любого из N элементов и при отказе более m ($0 < m < n$) из n параллельных элементов (резерв нагруженный). Считаем, что подключение резерва автоматическое и происходит достаточно быстро, т. е. временем его подключения можно пренебречь. Надежность автоматического коммутатора вновь не учитываем. Считаем также, что распределение времени безотказной работы каждого элемента последовательной группы подчиняется некоторому закону $F_j(t)$, $j=1, \dots, N$, а каждого элемента параллельной группы — закону $\Phi_i(t)$, $i=1, \dots, n$.

Для периодического наблюдения за состоянием элементов системы назначается некоторый интервал времени Δt и в моменты $t_i = i\Delta t$ ($i=0, 1, 2, \dots$) контролируется состояние всех элементов системы. Стоимость контроля, как и прежде, учитывать не будем.

Если в момент контроля обнаруживается отказ одного из N последовательно соединенных элементов или отказ $(m+1)$ -го из параллельных элементов (в любой последовательности), то система подвергается плановому аварийно-профилактическому ремонту. Во время проведения этого ремонта неотказавшие элементы системы проверяются в таком объеме (а отказавшие заменяются или капитально ремонтируются и устраняются все последствия имевших место отказов), что после ре-

монта система считается полностью обновленной. Аварийно-профилактический ремонт длится конечное время и стоимость его проведения равна $\tilde{c}_{а.п}$. Если при контроле выяснено, что не произошло отказа ни одного из последовательно работающих элементов, а число отказов параллельно работающих элементов не превосходит m , то осуществляется либо замена отказавших элементов (проводится плановая предупредительная профилактика системы, длящаяся конечное время, причем стоимость ее проведения есть $\tilde{c}_{п.п}$) и проверка неотказавших, либо принимается решение об эксплуатации системы в течение времени Δt до следующего момента контроля, и далее, описанная процедура принятия решения повторяется. Очевидно, $\tilde{c}_{а.п} > \tilde{c}_{п.п}$. Считаем, что момент $t_0 = 0$ есть момент последней регенерации.

Обозначим через X_i ($X_i = 0, 1, \dots, n$) число отказавших элементов в системе в момент наблюдения t_i ($i = 0, 1, 2, \dots$). Решение о контроле, о проведении профилактики или ремонта системы принимается в зависимости от значений X_i ($X_i = 0, 1, \dots, n$) и t_i ($i = 0, 1, 2, \dots$).

Как и в задаче, описанной выше, выбор стратегии обслуживания системы определяется разбиением области $A = \{X_i, t_i\}$ (в рассматриваемом случае вновь $t_i = t_i$) на непересекающиеся подобласти A_1, A_2, A_3 ($A_1 \cup A_2 \cup A_3 = A$), при попадании в которые точки $\{X_i, t_i\}$ принимается решение соответственно о проведении контроля, профилактики или ремонта.

Наша задача заключается по-прежнему в том, чтобы некоторым оптимальным отказом определить области A_1 и A_2 . Для решения задачи вновь воспользуемся критерием $M_{с_i}$. Отметим, что по аналогии с выражениями (4.43) и (4.44) для рассматриваемого случая можно записать

$$1 - \left[\prod_{j=1}^N \frac{\bar{F}_j(t_k)}{\bar{F}_j(t_{k-1})} \left\{ 1 - \prod_{i \neq i_1, i_2, \dots, i_q} \left[1 - \frac{\bar{\Phi}_i(t_k)}{\bar{\Phi}_i(t_{k-1})} \right] \right\} \right] \leq \leq \frac{\tilde{c}_{п.п}}{(\tilde{c}_{а.п} - \tilde{c}_{п.п}) (k-1)}, \quad (4.45)$$

где i_1, i_2, \dots, i_q — номера отказавших элементов параллельной группы к моменту t_{k-1} , причем $0 \leq q < m$, $\bar{F}(t) = 1 - F(t)$, $\bar{\Phi}(t) = 1 - \Phi(t)$.

Ясно, что все точки $\{X_{k-1}, t_{k-1}\}$, удовлетворяющие условию (4.45), принадлежат области A_1 .

Соединение элементов системы по ветвящейся схеме

Пусть элементы системы или ее отдельного блока соединены по ветвящейся схеме (рис. 4.9). Все элементы нагружены и равнонадежны. Число ветвящихся структур равно l .

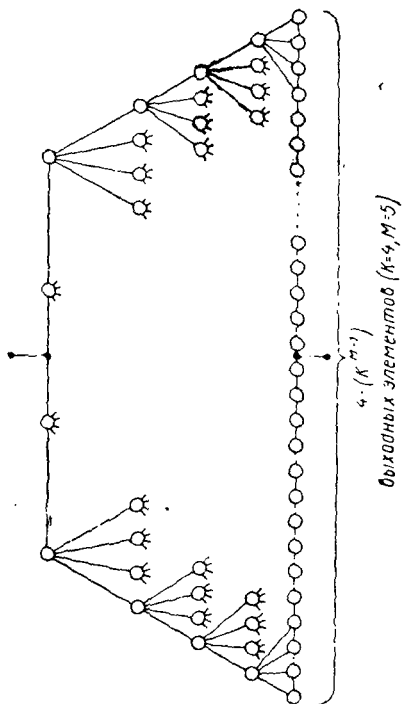


Рис. 4.9. Структура ветвящейся схемы соединения элементов системы

Так как все они одинаковы, то мы сначала рассмотрим одну ветвящуюся структуру (рис. 4.10), которую будем характеризовать «коэффициентом ветвления» K и количеством «рангов» M . Для определенности остановимся на структуре

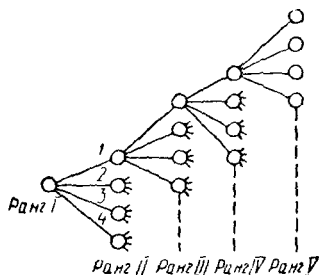


Рис. 4.10. Ветвящаяся структура

с $K=4$ и $M=5$. Считаем, что в структуре допустимое число отказов входных элементов равно m (всего на выходе структуры имеем K^{M-1} элементов, считаем также, что $m=q(K^{M-1})$, $0 < q < 1$).

Обозначим через $p(t)$ вероятность того, что любой из элементов структуры не откажет за время t . Тогда

$$q(t) = 1 - p(t).$$

Так как для рассматриваемой системы информацию о состоянии каждого ее элемента в дискретные моменты наблюдения $t_i = i\Delta t$ ($i=0, 1, 2, \dots$) использовать для принятия оптимальных решений на обслуживание системы [см. формулу (4.45)], не представляется возможным [здесь возникает необозримо большое число вариантов при определении условной вероятности — см. левую часть выражения (4.45)], то контроль состояния системы будем осуществлять путем наблюдения за состоянием ее выходных элементов. Стоимость контроля как и ранее учитывать не будем.

Если в момент контроля обнаружится отказ $(m+1)$ -го элемента на выходе ветвящейся структуры, то система подвергается плановому аварийно-профилактическому ремонту. Считаем, что после ремонта система обновляется. Стоимость проведения этого ремонта по-прежнему равна $\bar{c}_{a.п.}$.

Если при контроле выяснено, что число отказов выходных элементов не превосходит m , то осуществляется либо замена всех отказавших элементов (проводится плановая предупредительная профилактика системы, стоимость которой равна $\bar{c}_{ц.п.}$) и проверка неотказавших, либо принимается решение об эксплуатации системы в течение времени Δt до следующего момента контроля. Далее, описанная процедура принятия решения повторяется. Очевидно, что $\bar{c}_{a.п.} > \bar{c}_{ц.п.}$.

Считаем, что момент $t_0=0$ есть момент последней регенерации. Если через $X_i=0, 1, \dots, m+1$ обозначить число отказавших выходных элементов системы в момент наблюдения t_i ($i=0, 1, 2, \dots$), то решение о контроле, о проведении профилактики или ремонта системы принимается в зависимости от значений X_i, t_i .

Вновь выбор стратегии обслуживания системы определяется разбиением области $A = \{X_i, t_i\}$ на непересекающиеся подобласти A_1, A_2, A_3 ($A_1 \cup A_2 \cup A_3 = A$), при попадании в которые точки $\{X_i, t_i\}$ принимается соответствующее решение о проведении контроля, профилактики или ремонта системы.

Задача остается прежней: определить некоторым оптимальным образом области A_1 и A_2 . Для этой цели используем опять критерий Mc_i , однако, прежде чем в данном случае написать выражение, аналогичное (4.45), определим сначала функцию распределения времени до

отказа рассматриваемой системы (а значит, и ее надежность).

Рассмотрим сначала одну ветвящуюся структуру с произвольными величинами K и M и определим для нее вероятность того, что за время t на выходе откажут m элементов; обозначим эту вероятность через $P_m(K, M, t)$.

Выпишем рекуррентные соотношения для вероятностей $P_s(K, L, t)$ по L ($0 < s \leq K^L$). Предположим, что нам известна вероятность $P_s(K, L, t)$ при некотором фиксированном L ($1 \leq L \leq M-1$) для всех s . Определим вероятность $P_s(K, L+1, t)$:

$$P_s(K, L+1, t) = \sum_{n=0}^{\left[\frac{s}{K} \right]} P_n(K, L, t) C_{K^L \dots nK}^{s-nK} n^{s-nK}(t) q^{K^L-s}(t), \quad (4.46)$$

где $\left[\frac{s}{K} \right]$ - целая часть числа $\frac{s}{K}$, $0 \leq s \leq K^L$.

Очевидно, для $L=1$ (начальные условия):

$$P_0(K, 1, t) = n^*(t); \quad P_1(K, 1, t) = q(t); \quad 0 \leq s \leq K^L, \\ 1 \leq L \leq M-1. \quad (4.47)$$

Следовательно, вероятности $P_s(K, M, t)$ и $P_m(K, M, t)$ можно подсчитать рекуррентно.

Для рассматриваемой системы в целом (т. е. для четырех ветвящихся структур) вероятность иметь за время t не более $4m$ отказавших элементов на выходе запишется так:

$$P_{4m}(t) = \tilde{F}(t) = \sum_{s_1=0}^{4m} \sum_{s_2=0}^{4m-s_1} \sum_{s_3=0}^{4m-s_1-s_2} \sum_{s_4=0}^{4m-s_1-s_2-s_3} P_{s_1}(K, M, t) \times \\ \times P_{s_2}(K, M, t) P_{s_3}(K, M, t) P_{s_4}(K, M, t). \quad (4.48)$$

При этом, очевидно, области суммирования выбираются так, что $s_1 + s_2 + s_3 + s_4 < 4m$, т. е.

$$0 \leq s_4 \leq 4m - s_1 - s_2 - s_3; \quad 0 \leq s_3 \leq 4m - s_1 - s_2; \\ 0 \leq s_2 \leq 4m - s_1; \quad 0 \leq s_1 \leq 4m.$$

Теперь для системы запишем выражение, аналогичное (4.45):

$$1 - \frac{\bar{F}(t_k)}{F(t_{k-1})} \leq \frac{\tilde{c}_{n,n}}{(\tilde{c}_{n,n} - \tilde{c}_{t,n})(k-1)}, \quad (4.49)$$

где $\bar{F}(t)$ определяется зависимостью (4.48).

Алгоритмом (4.49) следует пользоваться при $X_{k-1} < 4m$. Все точки $\{X_{k-1}, t_{k-1}\}$, удовлетворяющие условию (4.49), принадлежат области A_1 .

4.5. ПЛАНИРОВАНИЕ ОБСЛУЖИВАНИЯ СИСТЕМЫ И ОТДЕЛЬНЫХ ЕЕ ЭЛЕМЕНТОВ ПРИ ХРАНЕНИИ

Определение плана проверок хранящейся системы при полном отсутствии информации о ее надежности. Случай конечного времени

Пусть новая система (блок, узел и т. п.), о надежности которой ничего не известно, должна храниться в течение заданного календарного времени T . Во время хранения система может отказаться. Для поддержания ее готовности к работе назначаются проверки. Если учесть стоимость проверок и потери от пребывания системы после отказа в неисправном состоянии, то ясно, что, с одной стороны, нельзя слишком часто назначать эти проверки, а с другой, — редкие проверки также приводят к большим потерям. (Вместо введенных здесь стоимостей проверок и затрат от пребывания системы в состоянии отказа можно рассматривать средние времена проверок и время пребывания системы в состоянии отказа. В этом случае решаемая ниже задача приводит к оптимизации плана проверок системы по готовности).

Считаем, что $F(t)$ — функция распределения времени до отказа ξ — неизвестна. Если в системе произошел отказ при хранении, то во время очередной проверки (при реализации выбранного плана проверок) он выявляется с вероятностью p . План проверок системы будем задавать последовательностью чисел

$$x_0 = 0 < x_1 < x_2 < \dots < x_n < x_{n+1} = T.$$

Пусть $G_{X_n}(\xi)$ является функцией потерь, которые имеют место, если отказ системы происходит в момент ξ и принят план проверок $X_n = (x_1, x_2, \dots, x_n)$. Так как си-

стема хранится в течение времени, не превышающего T , то будем считать $G_{X_n}(\xi) = G_{X_n}(T)$, если $\xi > T$.

Выражение для средних потерь за период хранения системы T будет иметь вид

$$M_F G_{X_n}(\xi) = \int_0^T G_{X_n}(u) dF(u) + G_{X_n}(T) [1 - F(T)]. \quad (4.50)$$

В связи с тем, что функция $F(t)$ неизвестна, будем определять такой план проверок X^* , при котором достигается минимума $\sup_F M_F G_{X_n}(\xi) = \mu_{X_n}$, т. е. необходимо определить

$$\mu_{X^*} = \min_{X_n} \mu_{X_n}. \quad (4.51)$$

Решим задачу для функции $G_{X_n}(\xi)$ следующего вида:

$$G_{X_n}(\xi) = \left\{ \begin{array}{l} nc, \xi > T; \\ nc + v(T - \xi), x_n < \xi \leq T; \\ kc + p[v(x_{k+1} - \xi) + c] + p \sum_{s=1}^{n-k-1} (1-p)^s \times \\ \quad \times [v(x_{k+s+1} - \xi) + (s+1)c] + \\ \quad + (1-p)^{n-k} [v(T - \xi) + (n-k)c], \\ x_k < \xi \leq x_{k+1} \quad (k=0, 1, \dots, n-1), \end{array} \right\} \quad (4.52)$$

где c — стоимость одной проверки;

v — потери в единицу времени от пребывания системы в состоянии отказа.

Преобразуем выражение для $G_{X_n}(\xi)$ при $x_k < \xi \leq x_{k+1}$, введя обозначения: $k+s=i$; $q=1-p$. Тогда получим

$$\begin{aligned} G_{X_n}(\xi) &= p(k+1)c + pv(x_{k+1} - \xi) + \\ &+ p \sum_{i=k+1}^{n-1} q^{i-k} [(i+1)c + v(x_{i+1} - \xi)] + q^{n-k} [nc + v(T - \xi)] = \\ &= g(x_1, x_2, \dots, x_n, \xi) \quad (k=0, 1, \dots, n-1). \end{aligned} \quad (4.53)$$

С учетом соотношения (4.53) перепишем выражение (4.52):

$$G_{X_n}(\xi) = \begin{cases} nc, & \xi > T; \\ nc + v(T - \xi), & x_n < \xi \leq T; \\ g(x_1, \dots, x_n; \xi), & x_k < \xi \leq x_{k+1} \quad (k=0, 1, \dots, n-1). \end{cases}$$

Средние потери $M_F G$ являются линейным функционалом относительно функции F . В силу следствия 1 (см. леммы 4.1 и 4.2), излагаемого ниже, $\sup_F M_F G$ достигается на вырожденной функции распределения со скачком в некоторой точке x . Для вырожденной функции распределения $F(t)M_F G = G_{X_n}(x)$, поэтому

$$\mu_{X_n} = \sup M_F G = \max_{0 < x \leq T} G_{X_n}(x).$$

Из выражения (4.52) следует, что

$$\begin{aligned} & \max_{0 < x \leq T} G_{X_n}(x) = \\ & = \max_k g(x_1, \dots, x_n; x_k + 0) = g(x_1, \dots, x_n; x_s + 0), \end{aligned}$$

где x_s — точка максимума.

Таким образом, максимальные средние потери достигаются на вырожденной функции распределения $F_{s,x}$ со скачком в точке x_s .

Найдем оптимальный план проверок X^* . Обозначим через C_n класс проверок объема n . Тогда для класса проверок \tilde{C}_M , объем которых не превышает M , имеем

$$\tilde{C}_M = \bigcup_{n < M} C_n.$$

Обозначим через C_n^* ($C_n^* \subset C_n$) класс проверок объема n , для которого x_1, \dots, x_n определяются из условия:

$$\left. \begin{aligned} g(x_1, \dots, x_n; x_0 + 0) &= g(x_1, \dots, x_n; x_1 + 0) = \\ &= \dots = g(x_1, \dots, x_n; x_n + 0); \\ \tilde{C}_{M^*} &= \bigcup_{n < M} C_n^*. \end{aligned} \right\} \quad (4.54)$$

Докажем, что оптимальный план проверок $X^* \in \tilde{C}_{N^*}$, т. е. число проверок в оптимальном плане не превосходит N^* , где N^* — максимальное положительное число, для которого выполняется неравенство

$$N(N-1) \leq \frac{2vT}{pc}. \quad (4.55)$$

Пусть X_n — произвольный план проверок из класса C_n , а $n > N^*$. Покажем сначала, что для любого плана проверок X_n найдется лучший план проверок X из класса $\tilde{C}_{N^*} : \mu_{X_n} \geq \mu_{X_{N^*}}$. Очевидно, что $\mu_{X_n} \geq nc$.

План проверок $X \in \tilde{C}_{N^*}$ возьмем из более узкого класса $C_{N^*}^*$, поэтому для такого плана имеем

$$\begin{aligned} g(x_1, \dots, x_{N^*}; x_0 + 0) &= g(x_1, \dots, x_{N^*}, x_1 + 0) \dots \\ &= \dots = g(x_1, \dots, x_{N^*}; x_{N^*} + 0), \end{aligned} \quad (4.56)$$

т. е. $X \in C_{N^*}^* \subset C_{N^*} \subset \tilde{C}_{N^*}$.

Покажем, что для выбранного плана

$$\mu_X = \left\{ vT + \frac{N^*c}{2} [(N^* + 1)p + 2] \right\} \frac{1}{pN^* + 1}. \quad (4.57)$$

В самом деле, при $k \leq N^* - 1$

$$\begin{aligned} g(x_1, \dots, x_{N^*}; x_k + 0) &= p(k+1)c + pv(x_{k+1} - x_k) + \\ &+ p \sum_{j=k+1}^{N^*-1} q^{i-k} [(i+1)c + v(x_{i+1} - x_k)] + \\ &+ q^{N^*-k} [N^*c + v(T - x_k)] = \\ &= c \left[p(k+1) + p \sum_{i=k+1}^{N^*-1} [(i+1)q^{i-k} + N^*q^{N^*-k}] \right] + \\ &+ v \left[p \sum_{i=k}^{N^*-1} q^{i-k} \sum_{s=k}^i \Delta S + q^{N^*-k} \sum_{s=k}^{N^*} \Delta S \right] = \\ &= c \left[p \sum_{i=k}^{N^*-1} (i+1)q^{i-k} + N^*q^{N^*-k} \right] + \\ &+ v \left[p \sum_{s=k}^{N^*-1} \Delta S \sum_{i=s}^{N^*-1} q^{i-k} + q^{N^*-k} \sum_{s=k}^{N^*} \Delta S \right] = \\ &= c \left[p \sum_{i=k}^{N^*-1} (i+1)q^{i-k} + N^*q^{N^*-k} \right] + v \sum_{s=k}^{N^*} q^{s-k} \Delta S; \\ &\Delta_k = x_{k+1} - x_k \quad (k = 1, 2, \dots, N^* - 1). \end{aligned} \quad (4.58)$$

Из выражения (4.58) следует, что для любого плана проверок из класса C_{N^*}

$$\left. \begin{aligned} g(x_1, \dots, x_{N^*}; x_{k-1} + 0) &= qg(x_1, \dots, x_{N^*}, x_k + 0) + \\ &+ cpk + v\Delta_{k-1}; \\ g(x_1, \dots, x_{N^*}; x_{N^*} + 0) &= cN^* + v\Delta_{N^*} \end{aligned} \right\} (4.59)$$

($k=1, \dots, N^*$).

Если $g(x_1, \dots, x_{N^*}; x_{k-1} + 0) = g(x_1, \dots, x_{N^*}; x_k + 0)$ ($k=1, \dots, N^*$), то имеем дело с планом проверок $X \in C_{N^*}^*$, для которого из выражений (4.59) получим

$$g(x_1, \dots, x_{N^*}; x_k + 0) = ck + \frac{v}{p} \Delta_{k-1} \quad (k=1, 2, \dots, N^*). \quad (4.60)$$

С учетом зависимостей (4.56) и (4.60) имеем N^*-1 уравнение

$$\Delta_k = \Delta_{k-1} - \frac{cp}{v} \quad (k=\overline{1, N^*-1}) \quad (4.61)$$

и уравнение

$$\Delta_k = \frac{\Delta_{k-1}}{p} \quad (k=N^*). \quad (4.62)$$

Из уравнений (4.61) и (4.62) с учетом нормирующего условия $\sum_{i=0}^{N^*} \Delta_i = T$ непосредственно следует:

$$\left. \begin{aligned} T &= \Delta_{N^*}(1 + pN^*) + \frac{cp}{v} \frac{N^*(N^*-1)}{2}; \\ \Delta_{N^*} &= \frac{T - cpN^*(N^*-1) \frac{1}{2v}}{pN^* + 1}. \end{aligned} \right\} (4.63)$$

Поэтому для плана $X \in C_{N^*}^*$ на основании выражений (4.59) и (4.63) имеем

$$\begin{aligned} \mu_X &= g(x_1, \dots, x_{N^*}; x_{N^*} + 0) = N^*c + v\Delta_{N^*} = \\ &= \frac{vT + \frac{1}{2} N^*c [p(N^* + 1) + 2]}{pN^* + 1}, \end{aligned}$$

что и убеждает нас в справедливости зависимости (4.57).

Теперь покажем, что $\mu_{X_n} \geq \mu_X$, $X \in \tilde{C}_{N^*}$.

Так как $\mu_{X_n} \geq nc$ и $n > N^*$, то с учетом выражения (4.55)

$$\mu_{X_n} - \mu_X \geq nc - \frac{vT + \frac{1}{2} N^* c [p(N^* + 1) + 2]}{pN^* + 1} \geq \frac{c}{pN^* + 1} > 0.$$

Значит, оптимальный план $X^* \in \tilde{C}_{N^*}$.

Далее покажем, что $X^* \in C_{n^*}$ (n^* — произвольное), $n^* \leq N^*$ (n^* соответствует плану X^*) и определим значение n^* .

Убедимся в том, что произвольный план объема n ($X_n \in C_n$) хуже плана $X \in C_{n^*}$, для которого выполняется условие (4.54). Для этого необходимо доказать, что

$$\max_k \{g(x_1, \dots, x_n; x_k + 0)\} \geq \frac{vT + \frac{1}{2} nc [p(n+1) + 2]}{pn + 1}. \quad (4.64)$$

Покажем справедливость условия (4.64) методом от противного. Пусть справедливо обратное утверждение:

$$\max_k \{g(x_1, \dots, x_n; x_k + 0)\} < \frac{vT + \frac{1}{2} nc [p(n+1) + 2]}{pn + 1}.$$

Это значит, что при любом k

$$g(x_1, \dots, x_n; x_k + 0) < \frac{vT + \frac{1}{2} nc [p(n+1) + 2]}{pn + 1}. \quad (4.65)$$

Тогда

$$\begin{aligned} \sum_{k=0}^n g(x_1, \dots, x_n; x_k + 0) &= S_n < \\ < \frac{vT + \frac{1}{2} nc [p(n+1) + 2]}{pn + 1} &= S_n^*. \end{aligned} \quad (4.66)$$

Из выражений (4.59), добавляя тождество $pg(x_0, x_1, \dots, x_n; x_0 + 0) = pg(x_0, x_1, \dots, x_n; x_0 - 0)$, получим систему уравнений:

$$\left. \begin{aligned} pg(x_1, \dots, x_n; x_k + 0) &= g(x_1, \dots, x_n; x_k + 0) - \\ &- g(x_1, \dots, x_n; x_{k-1} + 0) + v\Delta_{k-1} - ckp \quad (k=1, 2, \dots, n); \\ pg(x_0, x_1, \dots, x_n; x_0 + 0) &= pg(x_0, x_1, \dots, x_n; x_0 - 0). \end{aligned} \right\} (4.67)$$

Суммируя левые и правые части системы уравнений (4.67), будем иметь

$$pS_n = -(1-p)g(x_0, x_1, \dots, x_n; x_0 + 0) + \\ + g(x_1, \dots, x_n; x_n + 0) - v \sum_{k=0}^{n-1} \Delta_k + cp \frac{n(n+1)}{2}.$$

Откуда с учетом того, что $g(x_1, \dots, x_n; x_n + 0) = cn + v\Delta_n$, имеем

$$pS_n + qg^*(x_0, x_1, \dots, x_n; x_0 + 0) = vT + \frac{1}{2}nc[2 - p(n+1)]. \quad (4.68)$$

Так как правая часть выражения (4.68) не зависит от плана проверок, то уравнение (4.68) справедливо и для плана из класса C_n^* . Поэтому

$$pS_n^* + qg(x_0, x_1, \dots, x_n; x_0 + 0) = \\ = pS_n + qg(x_0, x_1, \dots, x_n; x_0 + 0).$$

В силу неравенства (4.66) $S_n^* > S_n$. Следовательно,

$$g(x_0, x_1, \dots, x_n; x_0 + 0) > g^*(x_0, x_1, \dots, x_n; x_0 + 0) = \\ = \left\{ vT + \frac{nc}{2} [p(n+1) + 2] \right\} \frac{1}{pn+1},$$

что противоречит принятому предположению (4.65). Значит, при любом n план из класса C_n^* лучше другого плана того же объема. Поэтому с учетом предыдущего оптимальный план X^* надо искать в классе $\tilde{C}_{N^*}^*$, т. е. можно записать

$$X^* \in \tilde{C}_{N^*}^* \subset \tilde{C}_{N^*}.$$

Теперь определим объем оптимального плана X^* , т. е. при каком $n^* \leq N^*$ меняет знак разность

$$\mu_{X_n^*} - \mu_{X_{n-1}^*} = \Delta \mu_n;$$

$$X_n^* \in C_n^*; \quad X_{n-1}^* \in C_{n-1}^*.$$

Ясно, что

$$\Delta \mu_n = \frac{vT + \frac{1}{2}nc[2 + p(n+1)]}{pn+1} - \frac{vT + \frac{1}{2}(n-1)c(2+np)}{(n-1)p+1}.$$

После преобразований имеем

$$\Delta \mu_n = \frac{n^2 + \frac{n(1+q)}{p} + \frac{2(c-vTp)}{cp^2}}{2(pn+1)[1+(n-1)p]} \quad (n > 0). \quad (4.69)$$

Так как знаменатель выражения (4.69) положителен, то для определения величины n^* , обеспечивающей минимум функции $\mu_{x_n^*}$ как функции n , необходимо исследовать числитель выражения (4.69), представляющий собой квадратичный трехчлен. Обозначим его через $f(n)$.

При анализе поведения функции $f(n)$ необходимо учесть два варианта: либо n^* определяется как наибольшее n , для которого $f(n) \leq 0$ при $c < vTp$, т. е. $cp^2n^2 + 2(cpn + c - vTp) - cp^2n \leq 0$; либо $n^* = 0$, если $c \geq vTp$.

Покажем, что $n^* \leq N^*$, для которого еще справедливо условие (1.55), т. е.

$$N^2 - N - \frac{2vT}{cp} = \varphi(N) \leq 0.$$

Определим теперь в интервале $[0, N^*]$ искомое значение $n^* = N$. Из выражения (4.69) имеем

$$f(n) = n^2 + \frac{n(1+q)}{p} + \frac{2(c-vTp)}{cp^2} = n^2 - n - \frac{2Tv}{cp} + \left[\frac{2}{p^2} + \frac{2n}{p} \right],$$

откуда ясно, что при $n = N = 0$

$$f(0) > \varphi(0) = -\frac{2vT}{cp}$$

и
$$\varphi'(N) = 2N - 1 < f'(n) = 2N - 1 + \frac{2}{p}.$$

Зависимости $\varphi(N)$ и $f(N)$ представлены графиками на рис. 4.11.

Запишем явное выражение для плана X^* . Из зависимости (4.61) после преобразований получим

$$\Delta_k = p\Delta_N + \frac{cP}{v} (N - 1 - k). \quad (4.70)$$

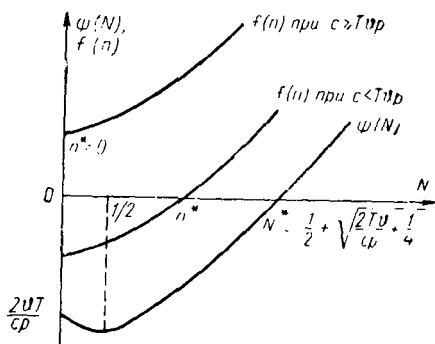


Рис. 4.11. К определению оптимального плана проверок СЛУ, находящейся на хранении

Подставив соотношение (4.63) в выражение (4.70), имеем

$$\begin{aligned} \Delta_k &= p \left\{ \frac{T}{pN+1} - (k+1) \frac{c}{v} + \frac{c [2N + 2N^2p - pN^2 + pN]}{2v(pN+1)} \right\} = \\ &= p \left\{ \frac{T}{pN+1} - (k+1) \frac{c}{v} + \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\}. \quad (4.71) \end{aligned}$$

Покажем, что для моментов проверок, определяющих оптимальный план X^* (см. также [54]), имеет место следующая формула при любом $k=i+1 > 0$, $k \leq N-1$:

$$x_i = ip \left\{ \frac{T}{pN+1} - \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} - \frac{i+1}{2} \frac{c}{v} \right\}. \quad (4.72)$$

Для этого достаточно показать, что разность Δ_k ($k=0, 1, \dots, n^*=N$), определенная с помощью выражения (4.72), совпадает с выражением (4.71):

$$\begin{aligned} x_{k+1} - x_k &= \Delta_k = kp \left\{ \frac{T}{pN+1} + \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\} + \\ &+ p \left\{ \frac{T}{pN+1} + \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\} - \frac{(k+1)p(k+2)c}{2v} \end{aligned}$$

$$\begin{aligned}
& -kp \left\{ \frac{T}{pN+1} + \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\} + \frac{kp(k+1)c}{2v} = \\
& = p \left\{ \frac{T}{pN+1} + \frac{2}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\} - \frac{cp}{2v} [2(k+1)] = \\
& = p \left\{ \frac{T}{pN+1} - (k+1) \frac{c}{v} + \frac{c}{2v} \frac{N[(N+1)p+2]}{pN+1} \right\}.
\end{aligned}$$

Определение плана проверок хранящейся системы при наличии минимальной информации о ее надежности. Конечное время

Рассмотрим вновь работу системы на конечном интервале времени. Будем считать, что отказ системы обнаруживается при первой проверке, следующей за моментом отказа ($p=1$), а стоимость проверок и стоимость пребывания системы в состоянии отказа примем, как и раньше, равной в соответствии с выражением (4.52), но в предположении, что $p=1$.

Пусть нам известно значение функции распределения $F(t)$ времени жизни системы ξ в некоторой точке τ , т. е.

$$F(\tau) = P\{\xi < \tau\} = \pi, \quad \tau \geq 0, \quad 0 < \pi < 1. \quad (4.73)$$

В силу того, что функция распределения $F(t)$, кроме одного значения, остается неизвестной, применим вновь минимаксный подход к решению задачи минимизации максимальной средней стоимости, связанной с проверками и отказами системы.

Будем искать план проверок системы X , который является набором $m+n$ точек, таких что

$$0 \leq x_1 \leq \dots \leq x_m \leq \tau \leq x_{m+1} \leq \dots \leq x_{m+n} \leq T.$$

Сформулируем задачу более строго. Поставим в соответствие любому плану проверок X функцию стоимости $G_X(\xi)$, которая является стоимостью, связанной с системой, имеющей время жизни ξ .

Тогда средняя ожидаемая стоимость для заданной функции распределения $F(t)$ есть вновь выражение вида (4.50).

Так как $F(t)$ неизвестна, за исключением точки τ , то рассмотрим, как и ранее, $\mu_X = \sup_F M_F G_X(\xi)$, где супремум берется по всем положительным функциям распределения случайных величин, удовлетворяющим соот-

пошению (4.73). Задача заключается в отыскании такого плана проверок $X=X^*$, при котором [см. выражение (4.51)] $\mu=\mu_{X^*}=\min \mu_X$, т. е. нужно найти план проверок, который минимизирует максимальные средние ожидаемые потери. Приведем конечный результат решения задачи [64].

Минимаксный план проверок системы $X^*=(x_1, \dots, x_m, \dots, x_{m+n})$ задается так. Для интервала справа от известной точки τ

$$x_{m+j}=\tau+j\left[\frac{T-j}{n+1}+\frac{c}{2v}\left(\frac{n(n+3)}{n+1}-j-1\right)\right] \quad (j=1, \dots, n), \quad (4.74)$$

где n является либо нулем, либо таким наибольшим положительным целым, при котором

$$n(n+1) \leq \frac{2v}{c}(T-\tau)-2. \quad (4.75)$$

Для интервала слева от точки τ

$$x_i=i\left[\frac{\tau}{m}+\frac{c}{2v}(m-i)\right] \quad (i=0, 1, \dots, m), \quad (4.76)$$

где m — такое наибольшее положительное целое, при котором

$$m(m-1) \leq \frac{2\pi v \tau}{(2-\pi)c}, \quad (4.77)$$

если

$$\frac{\pi}{1-\pi}\left[\frac{(k-m+1)c}{2}+v\left(\frac{x_{m+1}}{k+1}-\frac{\tau}{m}\right)\right]+(k-m)c \geq 0. \quad (4.78)$$

Здесь k равно наибольшему целому, при котором

$$k(k+1) \leq \frac{2\pi v x_{m+1}}{(2-\pi)c}. \quad (4.79)$$

В противном случае, если условие (4.78) не выполняется, план проверок, аналогичный выражению (4.76), задается следующим образом:

$$x_i=i\left[\frac{x_{m+1}}{m+1}+\frac{c}{2v}\right](m-i+1) \quad (i=0, 1, \dots, m) \quad (4.80)$$

при $m=k$, причем k определяется из условия (4.79).

При $n=0$ значения x_i задаются согласно выражению (4.76), если

$$\frac{\pi}{1-\pi} \left[\left(\frac{l(l+2)-1}{l+1} - m \right) \frac{c}{2} + v \left(\frac{T}{l+1} - \frac{\tau}{m} \right) \right] + (l-m)c \geq 0, \quad (4.81)$$

где l -- нуль или такое наибольшее положительное целое, при котором

$$l(l+1) \leq \frac{2\pi(vT-c)}{(2-\pi)c}. \quad (4.82)$$

Если условие (4.81) не выполняется, то x_i при $n=0$ задается как

$$x_i = i \left[\frac{T}{m+1} + \frac{c}{2v} \left(\frac{m(m+3)}{m+1} - i - 1 \right) \right] \quad (i=0, 1, \dots, m) \quad (4.83)$$

при $m=l$, определяемом из условия (4.82).

Минимаксная средняя ожидаемая стоимость μ определяется следующим образом:

$$\mu = \pi \left[\frac{(m+1)c}{2} + \frac{v\tau}{m} \right] + (1-\pi) \left[mc + \frac{n(n+3)}{2(n+1)}c + \frac{v(T-\tau)}{n+1} \right], \quad (4.84)$$

если $n=0$ и выполняется условие (4.81) или, если $n>0$ и выполняется условие (4.77);

$$\mu = \pi \left[\frac{l(l+3)}{2(l+1)}c + \frac{vT}{l-1} \right] + (1-\pi) [lc + v(T-\tau)], \quad (4.85)$$

если $n=0$ и условие (4.81) не выполняется;

$$\mu = \pi \left[\frac{k+2}{2}c + \frac{vx_{m+1}}{k+1} \right] + (1-\pi) \left[kc + \frac{n(n+3)}{2(n+1)}c + \frac{v(T-\tau)}{n+1} \right], \quad (4.86)$$

если $n=0$ и условие (4.77) не выполняется.

Как видно из приведенных результатов, в зависимости от значений стоимостей c , v и значения известной точки распределения π , τ/T минимаксный план будет принадлежать к двум типам планов.

Планом типа A назовем план с составляющими x_{m+1}, \dots, x_{m+n} , заданными согласно выражению (4.74), со значением n , которое определяется из условия (4.75), и с составляющими x_1, \dots, x_m , заданными согласно выра-

жению (4.76), со значением $m=r$, которое определяется из условия (4.78). Таким образом, план типа А включает в себя оптимальное планирование сначала на интервале (τ, T) , а затем на интервале $(0, \tau)$ при $x_m = \tau$.

Планом типа Б назовем план с составляющими x_{m+1}, \dots, x_{m+n} , заданными согласно выражению (4.74), со значением n , которое определяется из условия (4.75), и с составляющими x_1, \dots, x_m , заданными согласно выражению (4.80), со значением $m=k$, которое определяется из условия (4.79), если $n=0$, или из выражения (4.83) со значением $m=q$, которое определяется из условия (4.82), если $n=0$.

Таким образом, план типа Б включает в себя оптимальное планирование вновь сначала на интервале (τ, T) , а затем на интервале $(0, x_{m+1})$.

Если же $n=0$, то планирование осуществляется сразу на всем интервале $(0, T)$. Как и в случае плана типа А, так и в случае плана типа Б, план справа от точки τ — на интервале (τ, T) — будет оптимальным и таким, как в первом разделе п. 4.5.

В плане типа А проверка предусматривается в момент τ и интервал $(0, \tau)$ планируется оптимально согласно первому разделу п. 4.5. В плане типа Б не предусматривается проверка в момент τ , и интервал $(0, x_{m+1})$ или $(0, T)$, если $n=0$, будет планироваться так же, как и в первом разделе, но с незначительными изменениями. Сравнение моделей, рассматриваемых в первом и втором разделах, подробно изложено в работе [64].

Определение плана проверок хранящейся системы при наличии минимальной информации о ее надежности. Случай бесконечного времени

По-прежнему считаем, что отказы системы выявляются только при проведении проверок. Отличие решения задачи определения оптимального (в минимаксном смысле) плана проверок от решения, приведенного выше, будет заключаться только в том, что время работы системы примем равным бесконечности. Поэтому будем минимизировать максимальную среднюю стоимость, приходящуюся на единицу времени эксплуатации системы.

Наша стратегия будет заключаться в том, чтобы восстановить систему (до первоначального состояния) или заменить ее при первой проверке, во время которой об-

наруживается отказ (с вероятностью единица), или осуществить восстановление (замену) системы через время T , если за это время она проверялась, но при этих проверках не было обнаружено отказов. Замена через время T в последнем случае необходима в силу того, что задано, как и ранее, гарантированное значение

$$F(T) = P\{\xi < T\} = \pi_0, \quad T > 0, \quad (4.87)$$

где ξ — время безотказной работы системы (срок ее службы).

Здесь мы T трактуем как гарантийный ресурс работы системы и сверх времени T работа системы (без восстановления или замены) недопустима. Поэтому в соответствии с планом проверок X имеем вектор $X = (x_0, x_1, \dots, x_n, x_{n+1})$ со значениями, расположенными в интервале $(0, T)$ следующим образом:

$$0 = x_0 < x_1 < x_2 < \dots < x_n < x_{n+1} = T,$$

где $x_i (i=0, 1, \dots, n)$ — время, измеряемое от момента начала эксплуатации системы или от момента ее восстановления (замены).

Для любого заданного плана проверок X вновь введем в рассмотрение функцию стоимости $G_X(\xi)$, связанную с системой, которая имеет случайное время до отказа ξ . Обозначим через $R(\xi)$ отрезок времени, в течение которого система, имеющая случайный срок службы ξ , находится в эксплуатации, и назовем его циклом замены. Функционирование системы состоит из бесконечной последовательности таких циклов замен (при проверках и по истечении времени T).

Пусть C_i — стоимость, связанная с i -м циклом замены, а R_i — длина i -го цикла замены. Тогда средняя стоимость в единицу времени эксплуатации системы

$$\Phi(X) = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^{N(t)} C_i, \quad (4.88)$$

где $N(t)$ — число циклов, завершающихся к моменту t .

Предел (4.88) существует с вероятностью единица и может быть записан как

$$\Phi(X) = \frac{M_F C_i}{M_F R_i} \quad (4.89)$$

при условии, что $M C_i$ конечно.

Задача заключается в выборе плана X таким, чтобы минимизировать $\Phi(X)$, но так как $\Phi(X)$ зависит от функции распределения $F(t)$, которая неизвестна, за исключением единственного значения, то рассмотрим сначала

$$\mu^{(0)}(\underline{X}) = \max_F \Phi(X),$$

где максимум берется по всем функциям распределения положительных случайных величин, удовлетворяющих (4.87). В конечном итоге необходимо найти такой план проверок $X = X^*$, при котором

$$\mu^{(0)} = \min_X \mu^{(0)}(\underline{X}) = \mu^{(0)}(X).$$

Рассмотрим случай, когда стоимость, связанная с циклом замены, имеет вид (4.52). Приведем результат решения этой задачи, полученный в работе [65]:

$$(1-h)^{n+1} = 1 - \pi_0 - \frac{\pi_0 Q}{h(h-nQ)} [1 - (1-h)^n], \quad (4.90)$$

где $h = \mu/v$; $Q = c/(vT)$.

Из соотношения (4.90) при фиксированных Q и π_0 находится значение n^* , соответствующее оптимальному плану проверок. Далее план вычисляется рекуррентно из следующих соотношений:

$$\left. \begin{aligned} \pi_0 \gamma_{n^*} &= n^* Q - \pi_0 - h; \\ \pi_0 \gamma_{k-1} &= (1-h) \pi_0 \gamma_k - (1-\pi_0)(h - n^* Q) - \pi_0 k Q \\ &\quad (k = 1, 2, \dots, n^*), \end{aligned} \right\} \quad (4.91)$$

где $\gamma_i = x_i/T$.

Следует отметить, что уравнение (4.90) решается численными методами.

Некоторые математические обобщения и доказательства

При рассмотрении минимаксных критериев в задачах надежности первым этапом является определение наихудшей функции распределения $F(t)$ из множества допустимых функций, для которой средние потери $M_{FG_{X_n}}(\xi)$ при выбранном плане X_n были бы максимальны.

В рассмотренных выше и во многих других практических задачах средние потери имеют вид либо линейного, либо дробно-линейного функционала относительно функции $F(t)$. Поэтому для решения задачи определения наилучшей функции распределения представляет интерес описание вида функции распределения, на которой достигается экстремум исследуемого функционала. Ответ на этот вопрос дают доказываемые ниже утверждения.

Если $f(x)$, $g(x)$ — непрерывные ограниченные функции, $G(x)$ — функция распределения, то положим

$$I_G = \frac{\int f(x) dG(x)}{\int g(x) dG(x)}. \quad (4.92)$$

Заданы точки τ_1 ($\tau_1 < \tau_2 < \dots < \tau_m$) и ограничения

$$G(\tau_i) \in A_i, \quad 1 \leq i \leq m, \quad (4.93)$$

где A_i — любые замкнутые множества.

Требуется найти G , при котором $I_G = \max$.

Лемма 4.1. Максимум I_G в сформулированных условиях, если он существует, достигается при функциях G ступенчатого вида, имеющих не более одного скачка в каждом из интервалов $(-\infty, \tau_1)$, (τ_1, τ_2) , \dots , (τ_m, ∞) .

Введем уточнение. Пусть B_i — граница множества A_i .

Лемма 4.2. В условиях леммы 4.1 максимум I_G , если он существует, достигается на некоторой функции G ступенчатого вида, имеющей не более одного скачка на любом отрезке $[\tau_i, \tau_j]$, $0 \leq i < j \leq m+1$, таком, что для всех k , $i < k < j$, $G(\tau_k) \in B_k$.

Если говорить не совсем строго, то это означает следующее. Ограничение $G(\tau_k) \in A_k$ приводит к тому, что $G(\tau_k) \in B_k$, в противном случае оно как бы снимается.

Следствие 1. Пусть каждое множество $A_i = \{\pi_i\}$ состоит из одной точки $0 = \pi_0 \leq \pi_1 \leq \pi_2 \leq \dots \leq \pi_n \leq 1 = \pi_{n+1}$. Тогда максимум I_G , если он существует, достигается на функции $G(x)$ ступенчатого вида, имеющей на каждом из интервалов $(-\infty, \tau_1)$, \dots , (τ_n, ∞) скачок величины $\Delta\pi_i = \pi_{i+1} - \pi_i$.

Следствие 2. Утверждения лемм 4.1, 4.2 и следствия 1 справедливы для линейных функционалов вида

$$I_G = \int_{-\infty}^{+\infty} f(x) dG(x).$$

Рассмотрим непрерывный случай. Предположим, что решается задача максимизации функционала (4.92), причем функция распределения $G(x)$ подчинена ограничению

$$A_1(x) \leq G(x) \leq A_2(x), \quad (4.94)$$

где $A_1(x)$ и $A_2(x)$ — заданные неубывающие функции.

Можно попытаться искать решение задачи в классе функций $G(x)$ следующего вида: на одних отрезках $G(x) = A_1(x)$, а на других — $G(x) = A_2(x)$; в каждом интервале, где $A_1(x) < G(x) < A_2(x)$, функция $G(x)$ обладает не более чем одной точкой роста.

Справедливо следующее утверждение.

Лемма 4.3. Пусть Ω_G обозначает множество точек x , для которых либо $G(x) = A_1(x)$, либо $G(x) = A_2(x)$. Обозначим через L множество функций распределения $G(x)$, обращающих I_G в максимум при условии (4.94). Пусть M множество L , для каждого элемента G которого не существует $H \in L$, такого чтобы Ω_G было собственным подмножеством Ω_H . Тогда для любого $G \in M$ найдется $G^* \in M$, для которой будут выполнены следующие свойства:

- 1) $\Omega_{G^*} = \Omega_G$;
- 2) в любом интервале, где нет точек Ω_{G^*} , функция $G^*(x)$ имеет не более одной точки роста.

* *
*

Особенность рассмотренных выше задач заключается в том, что, отвлекаясь от инфинитезимального поведения реальных процессов и систем, мы интересовались только временными и отчасти качественными характеристиками. Такой подход во многих случаях приводит к необходимости исследования марковских и полумарковских моделей этих явлений.

Одна из рассмотренных проблем — обеспечение надежности адаптивных автопилотов; три другие — управление посадкой самолетов, управление дозаправкой самолетов горючим в воздухе и планирование обслуживания систем — типичны для класса организационных задач исследования операций. Решение этих проблем при проектировании позволит создать высокоэффективные комплексы.

Адаптивная система может обеспечить приемлемые характеристики управляемости для статически нейтрального или даже неустойчивого ЛА в широком диапазоне режимов полета и поэтому допускает более гибкий подход в выборе конкурирующих вариантов проектируемого объекта, чем САУ с программной коррекцией. Такой подход превратится в реальность, если отказ адаптивной системы управления станет столь же маловероятным событием, что и поломка планера ЛА.

Развитие в обозримом будущем высокоскоростных транспортных средств (многорежимных космических аппаратов, судов на подводных крыльях, аппаратов на воздушной подушке, экранопланов), предопределяет широкое использование высоконадежных адаптивных систем управления. Дальнейшее их развитие, по-видимому, все больше будет опираться на применение дискретных адаптивных алгоритмов, ЦВМ, систем с переменной структурой и нелинейной коррекцией.

СПИСОК ЛИТЕРАТУРЫ

1. **Александров А. Д.** Перспективные системы автоматического управления полетом. Информационные материалы Научного совета АН СССР по кибернетике, ВИНТИ АН СССР, 1970, № 7 (44), с. 20—28.

2. **Алексеев К. Б., Теряев Е. Д., Уколов И. С.** Реализация адаптивных алгоритмов управления с помощью БЦВМ. Информационные материалы Научного совета АН СССР по кибернетике, ВИНТИ АН СССР, 1970, № 7 (44), с. 7—19.

3. **Артамонов Г. Т., Михеев Ю. А.** Анализ системы разделения времени с однородными пользователями. — «Кибернетика», 1969, № 3, с. 94—97.

4. **Асратян Э. А., Симонов П. В.** Надежность мозга, М. АН СССР, 1963, с. 136.

5. **Барзилович Е. Ю., Каштанов В. А., Коваленко И. Н.** О минимаксных критериях в задачах надежности. — «Известия АН СССР. Техническая кибернетика», 1971, № 3, с. 87—93.

6. **Барзилович Е. Ю., Каштанов В. А.** Некоторые математические вопросы теории обслуживания сложных систем. М., «Советское радио», 1971, с. 271.

7. **Беляев Ю. К.** О числе выходов векторного случайного процесса за границу области. — «Теория вероятностей и ее применения», 1968, т. XII, № 2.

8. **Беляев Ю. К.** Лицевчатые марковские процессы и их приложение к задачам теории надежности. Труды VI Всесоюзного совещания по теории вероятностей и математической статистике. Госиздат Литовской ССР, Вильнюс, 1962, с. 309—323.

9. **Беляев Ю. К.** Новые результаты и обобщения задач типа пересечений. Дополнение к книге: Крамер Г., Лидбеттер М. Стационарные случайные процессы. М., «Мир», 1969, с. 398.

10. **Боднарчук В. Г.** и др. АНАЛИТИК (язык для описания вычислительных процессов с применением аналитических преобразований). Труды семинара «Теория автоматов и методы формализованного синтеза вычислительных машин и систем», изд. Научного совета по кибернетике АН УССР, Киев, 1968, с. 105.

— 11. **Боднер В. А.** Теория автоматического управления полетом. М., «Наука», 1964, с. 698.

12. **Большаков И. А.** Статистические проблемы выделения потока сигналов из шума. М., «Советское радио», 1969, с. 464.

13. Булинская Е. В. О среднем числе пересечений некоторого уровня стационарным гауссовским процессом. — «Теория вероятностей и ее применения», 1961, т. V, вып. 6, с. 474—478.

14. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. М., «Наука», 1966, с. 431.

15. Добрынин Е. Д., Курзенков Г. Д. Функциональная избыточность самонастраивающихся систем и ее влияние на надежность. Информационные материалы Научного совета АН СССР по кибернетике. ВИНТИ АН СССР, 1971, № 3 (50), с. 95—102.

—16. Дуб Дж. Л. Вероятностные процессы. М., ИЛ, 1956, с. 605.

17. Игнатьев М. Б., Москатов Г. К. О надежности одного класса самоорганизующихся систем с избыточностью. Труды IV симпозиума по проблеме избыточности в информационных системах, ч. 1, изд. Ленинградского института авиационного приборостроения и Научного совета АН СССР по кибернетике, Л., 1970, с. 273—283.

18. Калман Р. Вариационный принцип выбора оптимального фильтра из условия минимума квадратов ошибок. — В кн.: «Самонастраивающиеся автоматические системы», М., «Наука», 1964, с. 23—34.

19. Климов Г. П. Стохастические системы обслуживания. М., «Наука», 1966, с. 244.

20. Коваленко И. Н. Теория массового обслуживания. Итоги науки. Теория вероятностей. Математическая статистика. Теоретическая кибернетика. 1970, ВИНТИ АН СССР, 1971, с. 110.

21. Коваленко И. Н. Применение теории массового обслуживания к анализу и синтезу больших систем автоматического управления. — В кн.: «Современные методы проектирования систем автоматического управления», под ред. Б. Н. Петрова, В. В. Солодовникова и Ю. И. Гопчеева. М., «Машиностроение», 1967, с. 639—657.

22. Кокс Д. Р., Смит В. Л. Теория восстановления. М., «Советское радио», 1967, с. 299.

23. Кофман А., Крюон Р. Массовое обслуживание. Теория и приложения. М., «Мир», 1965, с. 302.

24. Кофман А. Методы и модели исследования операций. М., «Мир», 1966, с. 523.

25. Красовский А. А. Аналитическое конструирование контуров управления летательными аппаратами. М., «Машиностроение», 1969, с. 240.

26. Кузовков Н. Т. Динамика систем автоматического управления. М., «Машиностроение», 1969, с. 498.

27. Кухтенко В. И. Динамика самонастраивающихся систем со стабилизацией частотных характеристик. М., «Машиностроение», 1970, с. 232.

28. Мелен Д. Л. Применение самонастраивающейся системы для управления полетом. — В кн.: «Самонастраивающиеся автоматические системы». М., «Наука», 1964, с. 126—139.

29. Москатов Г. К. Надежность адаптивных систем управления полетом. Историко-технический анализ и взгляд в будущее. Труды VII Чтений К. Э. Циолковского. Изд. Института истории естествознания и техники АН СССР, 1973.

30. Москатов Г. К. Оценка надежности беспилотных самонастраивающихся систем управления. — В кн.: «Надежность и долговечность машин и приборов», под ред. Б. С. Сотскова, вып. 4, изд. НИИМАШ Министерства станкостроительной и инструментальной промышленности СССР, 1966, с. 153—164.

31. **Москатов Г. К.** Применение цепей Маркова для исследования надежности адаптивных автоматов стабилизации. Доклад на II симпозиуме по использованию избыточности в информационных системах. Материалы симпозиума. Изд. Ленинградского института авиационного приборостроения и Научного совета АН СССР по кибернетике, Л., 1966, с. 50—51.
32. **Москатов Г. К.** О линейчатом марковском процессе в двухконтурных самоприспосабливающихся системах. Доклад на III симпозиуме по проблеме избыточности в информационных системах. Материалы симпозиума. Изд. Ленинградского института авиационного приборостроения и Научного совета АН СССР по кибернетике, Л., 1968, с. 147—148.
33. **Москатов Г. К.** Применение принципа адаптации в целях обеспечения надежности систем стабилизации и управления. Информационные материалы Научного совета АН СССР по кибернетике, ВИНТИ АН СССР, 1971, № 3 (50), с. 186—205.
34. **Москатов Г. К.** Надежность адаптивных систем. М., «Советское радио», 1973, с. 104.
35. **Новиков О. А., Петухов С. И.** Прикладные вопросы теории массового обслуживания. М., «Советское радио», 1969, с. 399.
36. **Овчаров Л. А.** Прикладные задачи теории массового обслуживания. М., «Машиностроение», 1969, с. 324.
37. **Охоцимский Д. Е.** и др. Оптимальное управление при входе в атмосферу. — «Космические исследования», 1968, т. 6, № 1, с. 31—48.
38. **Петров Б. Н.** и др. Принципы построения и проектирования самонастраивающихся систем управления. М., «Машиностроение», 1972, с. 260.
39. **Петров Б. Н., Кухтенко А. И.** Инвариантные системы с переменными параметрами, средствами самонастройки и переменной структурой. — В кн.: «Современные методы проектирования систем автоматического управления», под ред. Б. Н. Петрова, В. В. Солодовникова и Ю. И. Топчеева, М., «Машиностроение», 1969, с. 13—78.
40. **Самонастраивающиеся системы.** Справочник под общей ред. П. И. Чинаева. Киев, «Наукова думка», 1969, с. 528.
41. **Саати Т.** Математические методы исследования операций. Пер. с англ. М., Восениздат, 1963, с. 423.
42. **Смит В.** Теория восстановления. Сб. переводов «Математика», 5:3, М., ИЛ, 1961.
43. **Сомик В. В., Коньков Н. Г.** Автомат проверяет самолет и ракету. М., Восениздат, 1967, с. 166.
44. **Столяров Г. К.** К вопросу об эффективности больших систем. — ДАН СССР, 1969, т. 189, № 2.
45. **Ушаков И. А., Коненков Ю. К.** Оценка эффективности функционирования сложных ветвящихся систем с учетом надежности. — В кн.: «Кибернетику на службу коммунизму», т. 2. М., «Энергия», 1964, с. 205—212.
46. **Хинчин А. Я.** Работы по математической теории массового обслуживания. М., Физматгиз, 1963, с. 236.
47. **Цыганков В.** и др. Системы заправки топливом в воздухе. — «Авиация и космонавтика», 1969, № 6, с. 40—41.
48. **Шерр А.** Анализ вычислительных систем с разделенным временем. М., «Мир», 1970, с. 135.
49. **Ширяев А. Н.** Статистический последовательный анализ. М., «Наука», 1969, с. 231.

50. **Электронная** вычислительная машина «МИР-2». Киев, «Реклама», 1971, с. 5.
51. **Barrer D. Y.** A waiting line problem characterized by impatient customers and indifferent clerks. *J. Oper. Res. Soc. of America*, 1955, vol. 3, pp. 360—366.
52. **Barrer D. Y.** Queuing with impatient customers and indifferent clerks. *J. Oper. Res. Soc. of America*, 1957, vol. 5, pp. 644—649.
53. **Blanchard B. S.** Cost effectiveness, system effectiveness, integrated logistics support and maintainability. *IEEE Trans. on Reliability*, vol. R—16, No 3.
54. **Derman C.** On minimax surveillance schedules. *Naval Research Logistics Quarterly*. 1961, vol. 8, No 4.
55. **Derman C., Sachs J.** Replacement of periodically inspected equipment. *Naval Research Logistics Quarterly*, 1960, vol. 7, No 4.
56. **Chow Y. S., Robbins H.** A martingale system theorem and applications. Proc. 4-th, Symp. on Math. Stat. and Probability, 1961. I, pp. 93—104.
57. **Cox D. R.** The analysis of non — markovain processes by the inclusion of supplementary variables. *Proc. Cambridge Philos. Soc.*, 1955, vol. 51, No 3, pp. 433—441.
58. **Holleman E. C., Adkins E. J.** Contribution of the X—15 Program to lifting entry technology. *J. of Aircraft*, 1964, vol. 1, No 6.
59. **Leadbetter M. R.** On crossings of arbitrary curves by certain Gaussian processes. *Proc. Amer. Math. Soc.*, 1965, vol. 16, No 1.
60. **Lindahl J. H., McGuire W. W.** Adaptive control flies the X—15. *Control Engineering*, 1962, vol. 9, No 10, pp. 93—97.
61. **Murray F. J., Miller K. S.** The mathematical basis for the error analysis of differential analyzers. *Mass. Inst. of Tech. Journal of Math. and Phys.*, 1953, No 2—3.
62. **Pearcey T., Bowen E. G.** Delays in the flow of air traffic. *J. of Royal Aeronaut. Soc.*, 1948, vol. 52, pp. 251—258.
63. **Pearcey T.** Delays in landing of air traffic. *J. of Royal Aeronaut. Soc.*, 1948, vol. 52, pp. 799—812.
64. **Roeloffs R.** Minimax surveillance schedules with partial information. *Naval Research Logistics Quarterly*, 1963, Dec., No 10.
65. **Roeloffs R.** Minimax surveillance schedules with replaceable units. *Naval Research Logistics Quarterly*, 1967, vol. 14, No 4.
66. **Semmelhack H. P., Hoppe S. G.** Learning adaptive flight control. Proc. 23 rd Nat. Electronics Conf., 1967, vol. 23, Chicago, pp. 2—7.
67. **Shen C. N., Cefola P. J.** Adaptive control for Mars entry based on sensitivity analysis. *AIAA Guidance, Control and Flight Dynamics Conf.*, AIAA Paper, No 68—835, Pasadena, Calif., 1968.
68. **Sutherland I. P.** Fly—by—wire flight control systems. Proc. Joint Symp. of Guidance, Flight Control and Flight Mechanics Panels of AGARD, Oslo, Sept., 1968.
69. **Thomson M. O., Welsh J. R.** Flight test experience with adaptive control systems. Proc. Joint Symp. of Guidance, Flight Control and Flight Mechanics Panels of AGARD, Oslo, Sept., 1968, pp. 38—59.
70. **Toll T. A., Fischel I.** The X—15 Project: Results and Research. *Astronautics and Aeronautics*, 1964. March.
71. **Kendall D. G.** Stochastic processes occuring in the theory of queues and their analysis by the method of imbedded Markov chains. *Ann. Math. Statist.*, 1953, vol. 24, pp. 338—354.

Предисловие	5
Глава I. Некоторые вопросы теории массового обслуживания. Оптимальная остановка процесса наблюдений	7
1.1. Предварительные замечания	7
1.2. Марковские процессы с конечным или счетным множеством состояний	16
1.3. Полумарковские процессы. Линейчатые процессы	29
1.4. Процессы восстановления	33
1.5. Оптимальное управление процессом восстановления	43
1.6. Правила остановки	48
Глава II. Надежность адаптивных систем управления полетом	60
2.1. Стабилизация и управление как процессы обслуживания «нетерпеливых клиентов»	60
2.2. Приспособляемость адаптивных систем к собственным частичным неисправностям	63
2.3. Структура входящего потока	73
2.4. Полумарковский процесс в адаптивных системах	76
2.5. Надежность контура адаптации	86
2.6. О структуре бортовых цифровых вычислительных машин	98
Глава III. Некоторые задачи управления движением на воздушных линиях	110
3.1. Управление посадкой групп самолетов	110
3.2. Управление заправкой самолетов в воздухе	120
Глава IV. Планирование обслуживания систем автоматического управления	123
4.1. Планирование замены отдельных элементов системы при мгновенной индикации отказов	123
4.2. Планирование замены элементов системы при отсутствии мгновенной индикации отказов	135
4.3. Планирование замены зарезервированных элементов системы в случае индикации отказов при проверках	139
4.4. Оптимальное обслуживание системы по наработке и состоянию составляющих элементов	143
4.5. Планирование обслуживания системы и отдельных ее элементов при хранении	154
Список литературы	172

*Игорь Николаевич Коваленко, Генрих Карлович Москатов,
Евгений Юрьевич Барзилович*

ПОЛУМАРКОВСКИЕ МОДЕЛИ В ЗАДАЧАХ ПРОЕКТИРОВАНИЯ
СИСТЕМ УПРАВЛЕНИЯ ЛЕТАТЕЛЬНЫМИ АППАРАТАМИ

Редактор издательства Л. В. Северьянов
Корректор В. Е. Блохина

Технич. редактор Л. Т. Зубко

Слано в набор 5/VII 1973 г. Подписано к печати 29/X 1973 г. Т-15937
Формат 84×108^{1/32} Бумага 2. Усл. печ. л. 9,24
Уч.-изд. л. 9,0 Тираж 3200 экз. Цена 55 коп. Изд. зак. 3656

Издательство «Машиностроение» Москва, Б-78, 1-й Басманный пер., 3.

Московская типография № 8 «Союзполиграфпрома»
при Государственном комитете Совета Министров СССР