АКАДЕМИЯ НАУК СССР

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

Том XV

1

МОСКВА · 1970

ОПТИМАЛЬНОСТЬ ОДНОРОДНОГО УПРАВЛЕНИЯ В ОДНОЙ ЗАДАЧЕ ТЕОРИИ НАДЕЖНОСТИ

В. А. МАЛЫШЕВ, А. И. ПИНСКИЙ

Система, функционирующая в интервале времени $[0, \infty]$, может находиться в двух состояниях: 0 (исправное), 1 (неисправное). В моменты t_i, $0 \leqslant t_1 < t_2 < \ldots < t_n < \ldots$ $\ldots < \infty$, $t_n \to \infty$, выбираемые заранее, делается мгновенная профилактика, т. е. в эти моменты система всегда находится в исправном состоянии. В промежутке между двумя соседними профилактиками $[t_i, t_{i+1}]$ время отказа системы распределено по закону $F(x)$, т. е. в промежутке $[t_i, x]$ система находится в состоянии 0, а в $[x, t_{i+1})$ в состоянии 1, причем $x - t_i$ распределено по закону $F(x)$. Будем считать, что известен также некоторый начальный момент профилактики $t_0 \leqslant 0$.

Рассмотрим две такие независимо работающие системы с одинаковыми распределениями отказов. Полученная система может находиться в одном из 4-х состояний: (0,0), (0,1), (1,0), (1,1). Будем считать неисправным состояние (1,1). На практике это соответствует случаю горячего резервирования. Пусть $\tau' = (t'_1, \ldots, t'_n \ldots)$ и $\tau'' = (t''_1, t''_2, \ldots)$ — выбираемые заранее моменты профилактик соответственно для первой и второй систем. Если мы платим за каждую профилактику и за время простоя обеих систем (время пребывания в состоянии (1,1)), то представляет интерес минимизация функционала

$$R(\tau', \tau'') = \overline{\lim_{T \to \infty}} \frac{1}{T} (C_1 n_T + C_2 M),$$

где n_T — число профилактик на интервале $[0, T]$; M — среднее время нахождения системы в состоянии (1,1) на интервале $(0, T]$; C_1 — цена одной профилактики, C_2 — цена единицы времени простоя. Мы будем интересоваться также оптимизацией стратегии на конечном отрезке, т. е. минимизацией функционала

$$R_T(\tau', \tau'') = C_1 n_T + C_2 M.$$

Пару (τ', τ'') назовем *управлением* (*стратегией*). Управление $(\tau', \tau'')^*$ оптимально, например, для $[0, T]$, если $R_T(\tau', \tau'')^* = \inf\limits_{(\tau',\tau'')} R_T(\tau', \tau'')$ и δ-оптимально, если $R_T(\tau',\tau'')^* \leqslant \inf R_T + \delta$. В работе находится оптимальное управление в некоторых предположениях на C_1, C_2, $F(x)$.

Замечание 1. Данная задача, во всяком случае, если $F(x) = 1 - e^{-x}$, может быть рассмотрена как задача управления марковским процессом по неполным данным. Пользуясь известным приемом (см., например, [1], стр. 155), можно свести ее к задаче управления по полным данным.

Именно, под состоянием системы будем понимать пару (P_1, P_2), где P_i — вероятность нахождения i-й системы в состоянии 0. Для простоты предположим, что время дискретно*, например, рассматривая моменты времени $n\Delta t$, $n = 1, 2, 3\ldots$, для достаточно малого Δt. Однородное управление тогда есть отображение конечного множества точек (P_1, P_2) в себя (конечность множества следует из ограниченности интервалов между профилактиками). Множество допустимых управлений также конечно, и функционал R является частным случаем функционала R_x^δ, рассматриваемого в работе [2]. Используя результаты этой работы можно доказать оптимальность однородного в смысле работы [1] управления в нашей задаче.

В дальнейшем, однако, мы понимаем однородное управление в гораздо более точном смысле. А именно, управление назовем *однородным*, если, во-первых, моменты профилактик первой и второй системы чередуются (такую стратегию будем называть *чередующейся*) и, во-вторых, интервалы между соседними профилактиками равны. Пусть $\widetilde{T}$ — класс чередующихся стратегий.

Лемма 1. $\inf\limits_{(\tau', \tau'')} R(\tau', \tau'') = \inf\limits_{(\tau', \tau'')\in\widetilde{T}} R(\tau', \tau'')$.

Доказательство. Если заменить некоторую стратегию на чередующуюся с профилактиками в те же моменты времени, то на каждом интервале времени между двумя соседними профилактиками среднее время пребывания системы в состоянии (1,1) может, как нетрудно заметить, лишь уменьшиться.

Далее будем рассматривать лишь чередующиеся стратегии (τ', τ''). Рассмотрим последовательность $\bar{x} = (x_1,\ldots, x_n\ldots)$ разностей между соседними профилактиками. Функционал R_T может быть записан тогда следующим образом:

$$R_T(\bar{x}) = C_1 n + C_2 \sum_{k=1}^{n-1} \int\limits_0^{x_{k+1}} F(t)\, F(t + x_k)\, dt + C_2 \int\limits_0^{T-\sum\limits_0^n x_i} F(t)\, F(t + x_n)\, dt.$$

Рассмотрим функционал R_T на отрезке $[0, T]$, $T = \sum\limits_{i=0}^{n-1} x_i$, в предположении $x_0 = x_n$, т. е. что последний отрезок между профилактиками равен времени работы до $t = 0$ одного из элементов.

Теорема 1. *Пусть $F(x)$ — трижды непрерывно дифференцируемая функция в окрестности U нуля. Тогда существует такое ε, зависящее только от $F(x)$, что при*

$$\max|x_k| \leqslant \varepsilon, \quad \frac{T}{n} = \frac{1}{m} \leqslant \varepsilon; \quad [0,2\varepsilon] \in U$$

имеет место

$$R_T(x) \geqslant R_T^{\cdot}\left(\frac{T}{n}, \ldots, \frac{T}{n}\right),$$

---

* Случай непрерывного времени, по-видимому, можно было бы рассмотреть, используя результаты работы [3].

т. е. однородное управление оптимально. При этом оценка ε может быть получена из следующего соотношения, связывающего первые три производные $F(x)$ на отрезке $[0,\ \varepsilon)$:

$$\frac{\lambda}{6m} - \frac{\lambda}{m} C(\varepsilon)\left(\frac{4}{m} + \frac{5}{3}\varepsilon\right) - \frac{\lambda}{m} B(\varepsilon)\left(\frac{5}{3m}\varepsilon + \frac{3}{m^2}\right) - \frac{3}{m}\lambda A(\varepsilon)\left(\varepsilon + \frac{1}{m}\right) -$$
$$- \frac{A(\varepsilon)\, C(\varepsilon)}{m^3} - \frac{C^2(\varepsilon)}{m^3} - \varepsilon\frac{C^2(\varepsilon)}{m^2}\cdot\frac{5}{6} - 4\frac{C(\varepsilon)\, B(\varepsilon)}{m^4} - 4\frac{B^2(\varepsilon)}{m^4} -$$
$$- \frac{C(\varepsilon)\, A(\varepsilon)}{m^4}\cdot\varepsilon\left(\frac{1}{m} + \frac{31}{24}\varepsilon\right) - \frac{B(\varepsilon)\, A(\varepsilon)}{m^2}\cdot\frac{15}{8}\varepsilon^2 - 5\frac{C(\varepsilon)\, B(\varepsilon)}{m^3}\varepsilon -$$
$$- \frac{13A(\varepsilon)\,\varepsilon^2}{10} - \frac{\lambda A(\varepsilon)}{4m^2} \geqslant 0, \qquad (1)$$

где

$$\lambda = F'(0), \qquad C(\varepsilon) = \max_{x\in[0,\ \varepsilon]} |u'(x)|, \qquad B(\varepsilon) = \max_{x\in[0,\ \varepsilon]} |u''(x)|,$$
$$A(\varepsilon) = \max_{x\in[0,2\varepsilon]} |u'''(x)|; \qquad u(x) = \lambda x - F(x).$$

Доказательство. Достаточно доказать, что

$$\Delta \equiv R_T(x_0, \ldots, x_n) - R_T\left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}\right) =$$
$$= \sum_{i=0}^{n-1} \int_0^{1/m+\varepsilon_{i+1}} F(t)\, F\left(t + \frac{1}{m} + \varepsilon_i\right) dt - \int_0^{1/m} F(t)\, F\left(t + \frac{1}{m}\right) dt \geqslant 0.$$

Положим $x_i = \frac{1}{m} + \varepsilon_i$. Тогда $-\frac{1}{m} \leqslant \varepsilon_i \leqslant \varepsilon - \frac{1}{m}$; $\sum_{i=0}^{n-1} \varepsilon_i = 0$. Пользуясь обозначением $F(x) = \lambda x - u(x)$, получим

$$\Delta = \sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} F(x)\, F\left(x + \frac{1}{m} + \varepsilon_i\right) dx +$$
$$+ \int_0^{1/m} F(x)\left[F\left(x + \frac{1}{m} + \varepsilon_i\right) - F(x)\right] dx = A_1 + A_2 + A_3 + A_4 + A_5,$$

где

$$A_1 = \sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} \lambda x\left(\lambda x + \frac{\lambda}{m} + \lambda\varepsilon_i\right) dx,$$

$$A_2 = -\sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} \lambda x\cdot u\left(x + \frac{1}{m} + \varepsilon_i\right) dx,$$

$$A_3 = -\sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} u(x)\left(\lambda x + \frac{\lambda}{m} + \lambda\varepsilon_i\right) dx,$$

$$A_4 = \sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} u(x)\, u\left(x + \frac{1}{m} + \varepsilon_i\right) dx,$$

$$A_5 = \sum_{i=0}^{n-1} \int_0^{1/m} \lambda x\left[u\left(x + \frac{1}{m}\right) - u\left(x + \frac{1}{m} + \varepsilon_i\right)\right] dx.$$

Исследуем сначала член A_1. После интегрирования имеем

$$A_1=\lambda^2\left[\frac{3}{2}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i^2+\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i\varepsilon_{i+1}+\frac{1}{3}\sum_{i=0}^{n-1}\varepsilon_i^3+\frac{1}{2}\sum_{i=0}^{n-1}\varepsilon_i\varepsilon_{i+1}^2\right].$$

Лемма 2. *Если* $\sum_{i=0}^{n-1}\varepsilon_i=0$ *и* $\varepsilon_i\geqslant-\frac{1}{m}$ *для всех* i, *то* $A_1\geqslant\frac{1}{6}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i^2$.

Доказательство. Преобразуем A_1 следующим образом:

$$A_1=\frac{1}{3}\sum_{i=0}^{n-1}\left[\frac{1}{m}\varepsilon_{i+1}^2+\varepsilon_i\varepsilon_{i+1}^2+\varepsilon_{i+1}^3+\frac{1}{m}\varepsilon_i\varepsilon_{i+1}\right]+$$

$$+\frac{1}{6}\sum_{i=0}^{n-1}\left(\frac{1}{m}\varepsilon_{i+1}^2+\varepsilon_i\varepsilon_{i+1}^2\right)+\frac{2}{3}\sum_{i=0}^{n-1}\frac{1}{m}\left(\frac{1}{2}\varepsilon_i^2+\varepsilon_i\varepsilon_{i+1}+\frac{1}{2}\varepsilon_{i+1}^2\right)+$$

$$+\frac{1}{3}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i^2=B_1+B_2+B_3+B_4.$$

Можно сделать следующие оценки:

$$B_2=\frac{1}{6}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_{i+1}^2\left(\frac{1}{m}+\varepsilon_i\right)\geqslant 0,\qquad B_3=\frac{2}{3}\frac{1}{m}\sum_{i=0}^{n-1}\left(\frac{\varepsilon_i}{\sqrt{2}}+\frac{\varepsilon_{i+1}}{\sqrt{2}}\right)^2\geqslant 0.$$

Докажем, что $B_1\geqslant-\frac{1}{6}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i^2$, откуда и будет следовать оценка леммы 2. Имеем

$$B_1=\frac{1}{3}\sum_{i=0}^{n-1}\varepsilon_{i+1}(\varepsilon_i+\varepsilon_{i+1})\left(\frac{1}{m}+\varepsilon_{i+1}\right)=\frac{1}{3}\sum_{i=0}^{n-1}r_i,$$

где

$$r_i=\varepsilon_{i+1}(\varepsilon_i+\varepsilon_{i+1})\left(\frac{1}{m}+\varepsilon_{i+1}\right).$$

В тех случаях, когда либо ε_i и ε_{i+1} имеют один знак, либо $|\varepsilon_i|\leqslant|\varepsilon_{i+1}|$, мы имеем $r_i\geqslant 0$. Рассмотрим два других случая: 1) $\varepsilon_i\geqslant 0$, $\varepsilon_{i+1}\leqslant 0$, $\varepsilon_i>|\varepsilon_{i+1}|$, т. е. $\varepsilon_{i+1}=-\delta\varepsilon_i$; $0\leqslant\delta\leqslant 1$. Тогда

$$r_i=-\frac{1}{3}\delta\varepsilon_i(\varepsilon_i-\delta\varepsilon_i)\left(\frac{1}{m}+\varepsilon_{i+1}\right)=-\frac{1}{3}\varepsilon_i^2\delta(1-\delta)\left(\frac{1}{m}+\varepsilon_{i+1}\right);$$

$r_i\geqslant(-1/12m)\varepsilon_i^2$, поскольку $0\leqslant\delta(1-\delta)\leqslant 1/4$ и $0\leqslant 1/m+\varepsilon_{i+1}\leqslant 1/m$.

2) $\varepsilon_i\leqslant 0$, $\varepsilon_{i+1}\geqslant 0$, $|\varepsilon_i|\geqslant\varepsilon_{i+1}$, т. е. $\varepsilon_{i+1}=-\delta\varepsilon_i$; $0\leqslant\delta\leqslant 1$. Тогда аналогичным образом

$$r_i=-\frac{1}{3}\delta\varepsilon_i(\varepsilon_i-\delta\varepsilon_i)\left(\frac{1}{m}+\varepsilon_{i+1}\right)\geqslant-\frac{1}{12}\frac{1}{m}\varepsilon_i^2.$$

Итак, $\sum_{i=0}^{n-1}r_i\geqslant-\frac{1}{6}\frac{1}{m}\sum_{i=0}^{n-1}\varepsilon_i^2$. Лемма доказана.

Основная идея оценки остальных членов состоит в том, что если в A_1 входят члены вида

$$\left(\frac{1}{m}\right)^{\alpha} \sum_{i=0}^{n-1} \varepsilon_i^{\beta} \varepsilon_{i+1}^{\gamma} \tag{2}$$

такие, что $\alpha + \beta + \gamma = 3$, $\beta + \gamma \geqslant 2$, то остальные члены A_i ($i \geqslant 2$) оцениваются членами вида (2) с $\alpha + \beta + \gamma \geqslant 4$ и $\beta + \gamma \geqslant 2$; подбором ε всегда можно добиться выполнения соотношения (1).

Оценим для примера один член

$$A_3 = -\sum_{i=0}^{n-1} \int_{1/m}^{1/m+\varepsilon_{i+1}} u(x)\left(\lambda x + \frac{\lambda}{m} + \lambda\varepsilon_i\right)dx = -\lambda \sum_{i=0}^{n-1} \int_0^{\varepsilon_{i+1}} u\left(y + \frac{1}{m}\right)\left(y + \frac{2}{m} + \varepsilon_i\right)dy -$$

$$- \sum_{i=0}^{n-1} \lambda \int_0^{\varepsilon_{i+1}} \left[u\left(\frac{1}{m}\right) + u'\left(\frac{1}{m}\right)y + u''\left(\frac{1}{m} + \theta y\right)\frac{y^2}{2}\right]\left(y + \frac{2}{m} + \varepsilon_i\right)dy.$$

Далее используются следующего типа оценки:

$$-\lambda \sum_{i=0}^{n-1} \int_0^{\varepsilon_{i+1}} u\left(\frac{1}{m}\right) y\,dy = -\lambda u\left(\frac{1}{m}\right)\frac{1}{2}\sum_{i=0}^{n-1} \varepsilon_{i+1}^2 \geqslant -\lambda u\left(\frac{1}{m}\right)\frac{1}{2}\sum_{i=0}^{n-1} \varepsilon_{i+1}^2;$$

$$-\lambda \sum_{i=0}^{n-1} \int_0^{\varepsilon_{i+1}} u'\left(\frac{1}{m}\right) y^2 dy = -\lambda u'\left(\frac{1}{m}\right)\frac{1}{3}\sum_{i=0}^{n-1} \varepsilon_{i+1}^3 \geqslant \lambda \left|u'\left(\frac{1}{m}\right)\right| \frac{1}{3}\varepsilon \sum_{i=0}^{n-1} \varepsilon_{i+1}^2.$$

С помощью аналогичной техники оцениваются остальные члены, откуда и следует оценка (1).

Замечание 2. Если $F(x) = \lambda x$ при $\lambda x \leqslant 1$ и $F(x) = 1$ при $\lambda x \geqslant 1$, то члены A_i, $i \geqslant 2$, отсутствуют при $x \leqslant 1/2$, и достаточным условием выполнения теоремы 1 является условие $\varepsilon \leqslant 1/\lambda$.

Следствие. Если не требовать, чтобы $x_0 = x_n$, то есть $t_{0,1}$ и $t_{0,2}$ могут быть какими угодно, то $(R_T^{\delta_{T/n}} - R_T^{\delta^*_T})/T = O(1/T)$, где δ^* — оптимальная стратегия на отрезке $[0, T]$, $\delta_{T/n}$ — управление с интервалом между соседними профилактиками T/n.

Остановимся вкратце на смысле применения теоремы 1. Практически важным является случай, когда C_1, C_2, $F(x)$ таковы, что невыгодно выбирать моменты профилактики таким образом, чтобы вероятность выхода из строя системы между двумя последовательными профилактиками была бы больше некоторого фиксированного числа, например, $p_0 = 0{,}01$. Иначе говоря, мы требуем, чтобы вероятность правильной работы системы между двумя профилактиками была бы близка к 1. Действительно, $p_0 \geqslant [F(\max_{1\leqslant k\leqslant n} x_k)]^2$ и, разрешая это неравенство относительно $\max_{1\leqslant k\leqslant n} x_k$, получаем

$$\max_{1\leqslant k\leqslant n} x_k \leqslant \varepsilon(p_0, F).$$

Возможен и альтернативный подход. Положим

$$u(\varepsilon, x_k, \gamma) = C_2 \int_0^{\varepsilon-\gamma} F(x)\,[F(t + x_k + \gamma) - F(t)]\,dt.$$

Если стратегию $\delta = (x_0, x_1, \ldots, x_k, \varepsilon, x_{k+2}, \ldots)$ заменить на
$$\delta' = (x_0, \ldots, x_k, \gamma, \varepsilon - \gamma, x_{k+2}, \ldots),$$

то для любого T такого, что $T \geqslant \sum_0^{k+2} x_i$, верно

$$R_T^{\delta_T} - R_T^{\delta'_T} \geqslant u(\varepsilon, x_k, \gamma) + C_1 \geqslant u(\varepsilon, 0, \gamma) + C_1.$$

Для любого ε существует γ_ε такое, что $u(\varepsilon, 0, \gamma_\varepsilon)$ достигает в γ_ε своего максимума. Если C_1 и C_2 таковы, что $u(\varepsilon, 0, \gamma_\varepsilon) \geqslant C_1$, то можно для выбора оптимальных управлений ограничиться классом стратегий $\Delta''(\varepsilon)$ таких, что $\sup\limits_i x_i \leqslant \varepsilon$ для $\delta = (x_0, \ldots) \in \Delta''(\varepsilon)$, и так как

$$u(\varepsilon, 0, \gamma_\varepsilon) = u(\varepsilon, 0, \varepsilon/2) = C_2 \int\limits_0^{\varepsilon/2} F(t)\left[F\left(t + \frac{\varepsilon}{2}\right) - F(t)\right] dt,$$

то условием принадлежности δ к $\Delta''(\varepsilon)$ является

$$C_2 \int\limits_0^{\varepsilon/2} F(t)\left[F\left(t + \frac{\varepsilon}{2}\right) - F(t)\right] dt \geqslant C_1.$$

Теорема 2. *Пусть* $\sup\limits_k x_k \leqslant \varepsilon$, *где* ε *и* $F(x)$ *удовлетворяют условиям теоремы* 1. *Тогда*

$$\inf_{\delta\in\Delta} \overline{\lim_{T\to\infty}} \frac{R_T^\delta}{T} = \lim_{T\to\infty} \frac{R_T^{\delta_{\theta_0}}}{T} = \frac{1}{\theta_0}\left[C_2 \int\limits_0^{\theta_0} F(t)\, F(t+\theta_0)\, dt + C_1\right],$$

где δ_{θ_0} — *стратегия* $\{\theta_0, \ldots, \theta_0\}$, *а* θ_0 *доставляет экстремум выражению*

$$\frac{1}{\theta}\left[C_2 \int\limits_0^{\theta} F(t)\, F(t+\theta)\, dt + C_1\right].$$

Этот минимум единствен, если $F'(t) > 0$, $F''(t) \leqslant 0$ *для всех* $t \geqslant 0$.

Авторы выражают благодарность Ю. К. Беляеву за постановку задачи и внимание к работе.

Поступила в редакцию
10.6.68

ЛИТЕРАТУРА

[1] А. Н. Ширяев, Некоторые новые результаты в теории управляемых случайных процессов. Trans. Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Prague, 1967, 131—203.

[2] О. В. Висков, А. Н. Ширяев, Об управлениях, приводящих к оптимальным стационарным режимам, Тр. Матем. ин-та им. В. А. Стеклова, 71 (1964), 35—45.

[3] В. Н. Рыков, Управляемые марковские процессы с конечными пространствами состояний и управлений, Теория вероят. и ее примен., XI, 2 (1966), 343—351.

OPTIMALITY OF HOMOGENEOUS CONTROL IN A PROBLEM OF RELIABILITY THEORY

V. A. MALYŠEV, A. I. PINSKIĬ (MOSCOW)

(Summary)

In this paper, the optimal prophylactic regime for two system working in the hot reserve is found.