Math-Net.Ru
Общероссийский математический портал

А. М. Горцев, Управление включением резервного прибора при косвенных наблюдениях за очередью,
Автомат. и телемех., 1977, выпуск 8, 25–30

https://www.mathnet.ru/at7416

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
https://www.mathnet.ru/rus/agreement

Параметры загрузки:
IP: 18.97.14.91
16 декабря 2025 г., 20:48:52

УДК 65.012.1.122:621.3.019.35

УПРАВЛЕНИЕ ВКЛЮЧЕНИЕМ РЕЗЕРВНОГО ПРИБОРА ПРИ КОСВЕННЫХ НАБЛЮДЕНИЯХ ЗА ОЧЕРЕДЬЮ

А. М. ГОРЦЕВ

(Томск)

Рассматриваются вопросы, связанные с включением резервного прибора в однолинейную систему массового обслуживания по косвенным наблюдениям за очередью. Подобные ситуации возникают тогда, когда непосредственно наблюдать очередь либо невозможно, либо очередь наблюдается с помехами, т. е. для целей управления включением резервного прибора наблюдается некоторая величина, стохастически связанная с длиной очереди.

1. Введение

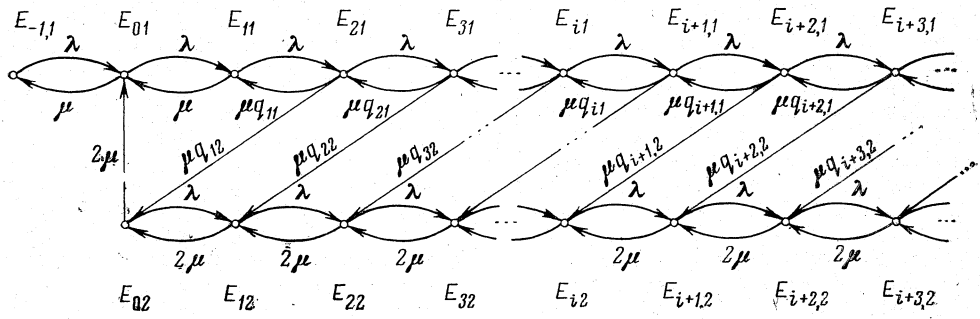
Задачам, связанным с управляемыми системами массового обслуживания (УСМО), посвящено много работ, на что указывает обзор, приведенный в [1]. Сравнительно небольшую часть составляют работы, в которых исследуются системы массового обслуживания (СМО) с резервными приборами (каналами), т. е. СМО с переменной структурой. Отметим работы, непосредственно связанные с вопросами, обсуждаемыми в данной статье. В [2—4] рассматривается система $M/M/2$ с одним основным и одним резервным каналами (симметричным и несимметричным); при достижении очередью длины n резервный прибор подключается к СМО, при достижении длины k ($k \leqslant n$) отключается. Задача оптимизации при этом заключается в определении уровней k и n, наилучших в смысле некоторого критерия оптимальности. В [5, 6] рассматриваются вопросы подключения к однолинейной СМО группы резервных приборов. Общим для задач, решаемых в указанных работах, является принцип управления по прямым наблюдениям за очередью, т. е. в каждый момент времени имеется информация о количестве заявок, ждущих обслуживания. Однако на практике могут возникать ситуации, при которых непосредственно наблюдать очередь либо невозможно, либо очередь наблюдается с помехами, т. е. для целей управления резервным прибором можно использовать лишь некоторую величину, в вероятностном смысле связанную с длиной очереди. Другими словами, наблюдаются значения некоторой стохастической функции $a(t)=\eta[i(t)]$, где $i(t)$ — длина очереди в момент времени t (t — текущее время). В частности, если $a(t)=i(t)+\xi(t)$, где $\xi(t)$ — аддитивный шум, имеют место наблюдения очереди с помехами.

Таким образом, в данном случае имеет место управление включением резервного прибора по косвенным наблюдениям за очередью. При этом задача оптимизации включения резервного прибора заключается в определении некоторого уровня $a^0=\beta$ (наилучшего в смысле некоторого критерия), превышение которого величиной $a(t)$ ($a(t) \geqslant \beta$) приводит к включению резервного прибора и к отключению при $a(t)<\beta$.

В настоящей статье при некоторых ограничениях на дисциплину обслуживания рассмотрим задачу определения оптимального уровня β включения резервного прибора.

2. Постановка задачи

Рассмотрим в простейших предпосылках (пуассоновский поток заявок с интенсивностью λ и экспоненциальное обслуживание с интенсивностью μ) однолинейную СМО с ожиданием, к которой может подключаться однотипный резервный прибор, причем загрузка системы $\rho=\lambda/\mu<2$. Для целей управления включением резервного прибора наблюдается некоторая неотрицательная величина a $(a\geqslant 0)$, в вероятностном смысле связанная с длиной очереди i, т. е. известно, например, условное распределение вероятностей $p(a/i)$. Обозначим через t_{k1}, t_{r2} $(k, r=1, 2, \ldots)$ моменты окончания обслуживания очередной заявки основным и резервным приборами соответственно. Ограничения на дисциплину обслуживания следующие:



1) если обслуживает только основной прибор, то в некоторый момент времени t_{k1} на обслуживание вызывается очередное требование, после чего наблюдается a и решается вопрос о включении резервного прибора; если $a\geqslant\beta$ $(\beta>0)$, то в момент времени t_{k1} резервный прибор подключается к основному; 2) если обслуживают оба прибора, основной и резервный, то в некоторые моменты времени t_{k1}, t_{r2} наблюдается значение a и решается вопрос об отключении одного из приборов; если $a=0$ (последнее означает отсутствие заявок в очереди) в некоторый момент времени t_{k1}, то отключается (т. е. становится резервным) основной прибор, если $a=0$ в некоторый момент времени t_{r2}, то отключается резервный прибор (здесь возможна замена основного прибора на резервный и наоборот — симметричность каналов обслуживания); в противном случае $(a>0)$ на обслуживание вызывается очередное требование.

Таким образом, на i-м шаге (т. е. когда в очереди на обслуживание находятся i требований) правило включения резервного прибора в силу стохастической связи наблюдаемой величины a с длиной очереди i является рандомизированным: резервный прибор подключается к основному с вероятностью q_{i2} и не подключается с вероятностью q_{i1} $(q_{i1}+q_{i2}=1)$, где $q_{i1}<1$ $(i\geqslant 1)$,

$$q_{i1}=q_{i1}(\beta)=\int_0^{\beta} p(a/i)\,da; \quad q_{i2}=q_{i2}(\beta)=\int_{\beta}^{\infty} p(a/i)\,da.$$

В силу дисциплины обслуживания $q_{01}=1$ (т. е. когда длина очереди равна нулю, то $a=0$, и с вероятностью единица выполняется неравенство $a<\beta$). При сделанных предположениях рассмотренная УСМО будет марковской. Обозначим через E_{i1} состояние системы, при котором в очереди находятся i требований и обслуживает только основной прибор $(i\geqslant 0)$, через E_{i2} — состояние, при котором в очереди находятся i требований и обслуживают как основной, так и резервный приборы $(i\geqslant 0)$. Кроме того, возможно еще одно состояние $E_{-1,1}$, когда длина очереди равна нулю и не обслуживают ни основной, ни резервный приборы. На рисунке изображен стохастический граф переходов, описывающий поведение приведенной УСМО, где каждой дуге графа поставлены в соответствие инфинитезимальные харак-

теристики (интенсивности переходов), причем петли в каждом состоянии опущены.

Введем в рассмотрение следующие характеристики исследуемой системы: $I(\beta, \rho)$ — средняя длина очереди на обслуживание, $Q(\beta, \rho)$ — вероятность работы резервного прибора. Обозначим через C_1 потери, которые несет СМО от ожидания одного требования в очереди в единицу времени, через C_2 — потери в связи с амортизацией резервного прибора в единицу времени ($C_1>0$, $C_2>0$). Тогда средние суммарные потери, которые несет система, запишутся в виде

$$L(\beta, \rho, C_1, C_2)=C_1 I(\beta, \rho)+C_2 Q(\beta, \rho). \tag{1}$$

Таким образом, задача оптимизации стратегии обслуживания резервным каналом при косвенных наблюдениях за очередью заключается в минимизации (1) по β при фиксированных ρ, C_1, C_2. Отметим, что аналогичный критерий принят в [6].

3. Определение стационарных вероятностей и характеристик системы

Обозначим через $P(i, j)$ стационарную вероятность того, что в очереди находятся i требований и обслуживают j приборов ($i\geqslant 0$, $j=1, 2$). $P(-1, 1)$ — стационарная вероятность состояния $E_{-1,1}$. Тогда для стационарных вероятностей справедливы соотношения

$$P(0, 1)=P(-1, 1)\rho,$$

$$P(i, 1)=P(-1, 1)\rho^{i+1}\prod_{j=1}^{i}\alpha_j, \quad i\geqslant 1; \tag{2}$$

$$P(0, 2)={}^1/_2 P(-1, 1))\rho^2(1-\alpha_1),$$

$$P(1, 2)=\frac{1}{2}P(-1, 1)\rho^3\left\{\frac{1}{2}(1-\alpha_1)+\sum_{j=2}^{\infty}\rho^{j-1}q_{j2}\prod_{k=2}^{j+1}\alpha_k\right\}, \tag{3}$$

$$P(i, 2)=P(-1, 1)\rho^2\left\{\frac{1}{2}(\rho/2)^i(1-\alpha_1)+(2-\rho)^{-1}\left[(1-\right.\right.$$

$$-(\rho/2)^i)\sum_{j=i+1}\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k-(\rho/2)^i\sum_{j=2}^{i}[1-$$

$$\left.\left.-(2/\rho)^{j-1}]\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k\right]\right\}, \quad i\geqslant 2;$$

$$P(-1, 1)=\left\{\frac{2+\rho}{2-\rho}+\frac{1-\rho}{2-\rho}\rho^2\alpha_1+\rho\sum_{j=2}^{\infty}(1+(\rho/2)q_{j2}\alpha_{j+1})\times\right. \tag{4}$$

$$\times\rho^j\prod_{k=1}^{j}\alpha_k+\rho^2(2-\rho)^{-1}\sum_{i=2}^{\infty}\left[(1-(\rho-2)^i)\sum_{j=i+1}^{\infty}\rho^j q_{j2}\times\right.$$

$$\left.\left.\times\prod_{k=1}^{j+1}\alpha_k-(\rho/2)^i\sum_{j=2}^{i}[1-(2/\rho)^{j-1}]\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k\right]\right\}^{-1},$$

где $\alpha_j = \left(1+\rho - \frac{\rho q_{j1}}{1+\rho} - \frac{\rho q_{j+1,1}}{1+\rho} - \ldots - \frac{\rho q_{r1}}{1+\rho} - \ldots\right)^{-1}$ $(j=1, 2, \ldots)$ — бесконечная цепная дробь [7]. Вывод формул (2)—(4) дан в приложении.

Формулы (2)—(4) позволяют получить интересующие нас характеристики системы

$$
\begin{aligned}
(5)\qquad I(\beta,\rho) = P(-1,1)\rho\Big\{&\rho\alpha_1+\rho^2(2-\rho)^{-2}(1-\alpha_1)+\\
&+\sum_{i=2}^{\infty}[i+(\rho/2)q_{i2}\alpha_{i+1}]\rho^i\prod_{k=1}^{i}\alpha_k+\\
&+\rho(2-\rho)^{-1}\sum_{i=2}^{\infty} i\Big[(1-(\rho/2)^i)\sum_{j=i+1}^{\infty}\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k-\\
&-(\rho/2)^i\sum_{j=2}^{i}[1-(2/\rho)^{j-1}]\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k\Big]\Big\},
\end{aligned}
$$

$$
\begin{aligned}
(6)\qquad Q(\beta,\rho) = P(-1,1)\rho^2\Big\{&(2-\rho)^{-1}(1-\alpha_1)+{}^1/_2\sum_{j=2}^{\infty}\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k+\\
&+(2-\rho)^{-1}\sum_{i=2}^{\infty}\Big[(1-(\rho/2)^i)\sum_{j=i+1}^{\infty}\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k-\\
&-(\rho/2)^i\sum_{j=2}^{i}[1-(2/\rho)^{j-1}]\rho^j q_{j2}\prod_{k=1}^{j+1}\alpha_k\Big]\Big\}.
\end{aligned}
$$

Подставляя (5) и (6) в (1), получаем выражение для средних суммарных потерь (1).

Очевидно, что полученная задача оптимизации ввиду того, что в выражения (5) и (6) входят бесконечные цепные дроби α_j $(j\geqslant 1)$, даже в простейших случаях (простота определяется заданием распределения $p(a/i)$, например $p(a/i)$ — равномерное распределение) допускает лишь численное решение. Принципиальных трудностей при получении численного решения нет (так как функция (1) — функция одного переменного β, $\beta>0$), последнее при фиксированных ρ, C_1, C_2 можно, например, получить прямыми методами поиска экстремума [8].

Отметим, что из (2)—(4) вытекают частные случаи: 1) $q_{i1}=q$ $(i\geqslant 1)$, что соответствует рандомизированному правилу включения резервного прибора, т. е. резервный прибор подключается к основному с постоянной вероятностью $p=1-q$ на каждом шаге (после каждого окончания обслуживания заявки основным прибором); 2) $p(a/i)=\delta(a-i)$ $(i\geqslant 1)$, последнее означает, что управление по косвенным наблюдениям вырождается в управление по прямым наблюдениям за очередью: резервный прибор подключается к основному при достижении очередью длины $i\geqslant n$ $(n\geqslant 2)$.

В заключение рассмотрим возможность других дисциплин обслуживания по косвенным наблюдениям за очередью. В рамках изложенной стратегии обслуживания существенным является момент наблюдения величины a. Если наблюдение величины a (в момент окончания обслуживания очередной заявки, когда работает один основной прибор) производить до вызова требования на обслуживание, то это может привести к «холостому ходу» резервного прибора. Эта возможность может реализоваться в случае, когда в очереди находится одна заявка. С другой стороны, если обслуживают оба прибора, основной и резервный, то наблюдение величины

a после вызова требования на обслуживание может привести к последующему отказу от обслуживания уже вызванного требования. Эта возможность, как и в первом случае, реализуется, когда в очереди находится одна заявка.

Область применения полученных результатов, как отмечалось в [6],— вычислительные комплексы, объединяющие в себе две и более ЭВМ. Обычно пользователь указывает примерное время решения задачи на ЭВМ (a_j). Оператор при решении задач на ЭВМ может выносить решение о подключении резервной ЭВМ (последняя в режиме нормальной загрузки используется лишь для хранения информации) на основании суммарного предполагаемого времени решения ($a=\Sigma a_j$) всех задач, стоящих в очереди, количество которых он может и не знать (например, когда задачи для решения поступают в пакетах). При этом оценку распределения $p(a/1)$ нетрудно сделать экспериментально (оценки распределений $p(a/i)$, учитывая независимость задач, поступающих на вычислительный центр, находятся на основании оценки $p(a/1)$), причем очевидно, что $a=0$ тогда и только тогда, когда задач для решения нет.

ПРИЛОЖЕНИЕ

Вывод формул (2) — (4). Для стационарных вероятностей $P(i, j)$, как нетрудно убедиться, имеют место прямые уравнения Колмогорова

$$\rho P(i, 1)-(1+\rho)P(i+1, 1)+q_{i+1,1}P(i+2, 1)=0 \quad (i\geqslant 0), \tag{П.1}$$

$$\rho P(i, 2)-(2+\rho)P(i+1, 2)+2P(i+2, 2)=-q_{i+2,2}P(i+3, 1) \quad (i\geqslant 0). \tag{П.2}$$

Рассмотрим уравнения (П.1). Непосредственной подстановкой выражения

$$P(i, 1)=P(0, 1)\rho^i\prod_{j=1}^{i}\alpha_j \quad (i\geqslant 1) \tag{П.3}$$

в (П.1) нетрудно показать, что (П.3) удовлетворяет (П.1). Отметим, что α_j является безусловно сходящейся бесконечной цепной дробью [7]. Действительно, обозначим

$$\alpha_{jr}=\left(1+\rho-\frac{\rho q_{j1}}{1+\rho}-\frac{\rho q_{j+1,1}}{1+\rho}-\ldots-\frac{\rho q_{j+r,1}}{1+\rho}\right)^{-1} \quad (r\geqslant 0),$$

т. е. α_{jr} — конечная цепная дробь. Тогда очевидно, что

$$(1+\rho)^{-1}<\alpha_{j0}<\alpha_{j1}<\ldots<\alpha_{jr}<\ldots<\alpha_j. \tag{П.4}$$

Таким образом, набор α_{jr} ($r\geqslant 0$) образует монотонно возрастающую последовательность. Кроме того (так как $q_{j1}<1$, $j\geqslant 1$), выполняется неравенство

$$\alpha_j\leqslant\left(1+\rho-\frac{\rho q^*}{1+\rho}-\frac{\rho q^*}{1+\rho}-\ldots-\frac{\rho q^*}{1+\rho}-\ldots\right)^{-1}=\alpha^*, \tag{П.5}$$

где $q^*=\sup q_{j1}$ ($j\geqslant 1$). Нетрудно показать, что

$$\alpha^*=(1+\rho-\sqrt{(1+\rho)^2-4\rho q^*})(2\rho q^*)^{-1}>0, \tag{П.6}$$

т. е. из (П.4), (П.5), (П.6) и произвольности j ($j\geqslant 1$) вытекает безусловная сходимость бесконечной цепной дроби α_j. Для стационарных вероятностей $P(-1, 1)$ и $P(0, 1)$ справедливо уравнение $\lambda P(-1, 1)=\mu P(0, 1)$, откуда следует первое равенство в (2), подстановка которого в (П.3) приводит ко второму равенству в формулах (2).

Рассмотрим уравнения (П.2). Решение разностного неоднородного уравнения (П.2) есть сумма [9] общего решения однородного уравнения и частного решения неоднородного (последние нетрудно получить)

$$P(0, 2)=R_1+R_2,$$

$$P(1, 2)=R_1+R_2(\rho/2), \tag{П.7}$$

$$P(i, 2)=R_1+R_2(\rho/2)^i+(2-\rho)^{-1}\sum_{j=2}^{i}[(\rho/2)^{i-j+1}-1]q_{j2}P(j+1, 1) \quad (i\geqslant 2),$$

где R_1 и R_2 — некоторые константы, определяемые из граничных условий, одно из которых, очевидно, есть $\lim P(i, 2)=0$ при $i\to\infty$. Нетрудно показать, что для выполнения этого условия необходимо, чтобы

$$R_1=(2-\rho)^{-1}\sum_{j=2}^{\infty} q_{j2}P(j+1, 1). \tag{П.8}$$

Константа R_2 легко выражается через $P(-1, 1)$ и R_1:

$$R_2=(\rho^2/2)(1-\alpha_1)P(-1, 1)-R_1. \tag{П.9}$$

Подставляя (П.8) и (П.9) в (П.7), учитывая при этом (2) и проделывая несложные преобразования, получаем (3). Остается показать, что для полученных стационарных вероятностей выполняется условие нормировки. Для этого необходимо установить абсолютную сходимость рядов $\sum_{i=-1}^{\infty} P(i, 1)$ и $\sum_{i=0}^{\infty} P(i, 2)$. Учитывая, что в силу (П.5) и (П.6) справедливо $P(i, 1)\leqslant P(0, 1)\gamma^i (i\geqslant 0)$, где $\gamma=\rho\alpha^*$ ($0<\gamma<1$ для $0<\rho<2$ и $0<q^*<1$), и что $P(i, 2)$ выражается через $P(i, 1)$ (см. (П.7)), нетрудно показать, что упомянутые ряды мажорируются сходящимися геометрическими прогрессиями. Вследствие этого условие нормировки имеет место, используя его, получаем (4).

Поступила в редакцию
11 октября 1976 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Рыков В. В.* Управляемые системы массового обслуживания. Сб. ВИНИТИ, Итоги науки и техники, серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика», т. 12, стр. 43—153. Изд-во ВИНИТИ, 1975.
2. *Bartoszewicz I., Rolski T.* Queueing systems with a reserve service channel. Zastosow. mat., vol. 11, No. 4, pp. 439—449, 1970.
3. *Singh V. P.* Queue-dependent servers. J. Eng. Math., vol. 7, No. 2, pp. 123—126, 1973.
4. *Горцев А. М.* Система массового обслуживания с резервным каналом. Техническая кибернетика, № 6, стр. 85, 1975.
5. *Афанасьева А. Г.* Система с включением резервного прибора. Техническая кибернетика, № 2, стр. 34—40, 1971.
6. *Горцев А. М.* Оптимальное управление включением группы резервных приборов в однолинейной системе массового обслуживания. Автоматика и телемеханика, № 12, стр. 78—87, 1975.
7. *Данилов В. А., Иванова А. Н., Исакова Е. К., Люстерник Л. А., Салехов Г. С., Хованский А. Н., Цлаф Л. Я., Янпольский А. Р.* Математический анализ. Физматгиз, 1961.
8. *Уайлд Д.* Методика поиска экстремума. «Наука», 1967.
9. *Гельфонд А. О.* Исчисление конечных разностей. «Наука», 1967.

CONTROL OF CONNECTING A STAND-BY INSTRUMENT IN INDIRECT OBSERVATIONS OF A QUEUE

A. M. GORTSEV

The paper is concerned with connection of a stand-by instrument into a single-line servicing system on the knowledge of indirect observations of the queue. This is the case where direct observation is either impossible or if the queue is observed with noises; in other words, for control of connecting a stand-by instrument a certain quantity stochastically related to the queue length is observed.