УДК 519.718

И.В. Павлов

ВЫЧИСЛЕНИЕ НЕКОТОРЫХ ПОКАЗАТЕЛЕЙ КАЧЕСТВА И НАДЕЖНОСТИ ДЛЯ СИСТЕМЫ С ПАРАЛЛЕЛЬНО НАГРУЖЕННЫМИ ЭЛЕМЕНТАМИ

Для модели системы с параллельно нагруженными элементами получены расчетные формулы для вычисления основных показателей качества и надежности ее функционирования: среднего числа исправно работающих элементов в системе, средних затрат в единицу времени на восстановление отказавших элементов системы в стационарном режиме и др.

E-mail: ipavlov@bmstu.ru

Ключевые слова: *надежность, интенсивность отказов, восстановление, подсистемы.*

Пусть имеется система, включающая в себя N основных элементов, работающих одновременно в параллельном нагруженном режиме. В процессе работы каждый из элементов может отказывать с интенсивностью отказов $\lambda(t) \equiv \lambda$, а наработка на отказ каждого отдельного элемента имеет экспоненциальное распределение с функцией надежности $P(t) = e^{-\lambda t}$ [1, 2]. Предполагается, что эффективность работы системы пропорциональна числу исправно работающих ее элементов и отказы различных элементов происходят независимо друг от друга. При этом отдельные элементы системы могут быть объединены в некоторые функциональные блоки (подсистемы), которые также могут отказывать. Отказавшие в процессе работы элементы или блоки (подсистемы) в дальнейшем способны восстанавливаться (заменяться новыми идентичными элементами или блоками). Аналогичные по смыслу модели рассматривались в работах [3—6] и др.

Основные показатели качества и надежности функционирования системы. Пусть в системе имеется L блоков (подсистем), каждый из которых включает в себя m элементов. Обозначим через $k_i(t)$ число отказавших модулей в i-м блоке к моменту времени t. Величину $k_i(t)$ будем называть состоянием i-го блока в момент времени t, предполагая далее, что элементы в различных блоках и внутри каждого отдельного блока отказывают независимо друг от друга. Тогда состояние $k_i(t)$ каждого i-го блока к моменту времени t имеет биномиальное распределение:

$$p_t(k) = P\{k_i(t) = k\} = C_m^k (1 - e^{-\lambda t})^k (e^{-\lambda t})^{m-k}, \quad (1)$$

где $p_t(k)$ — вероятность того, что отдельный блок в момент времени t находится в состоянии k, т. е. содержит в своем составе k отказавших элементов; $k = 0, 1, ..., m$.

Обозначим через $N_k = N_k(t)$ число блоков, таких, что $k_i(t) = k$, т. е. имеющих в своем составе ровно k отказавших элементов в момент времени t. Вектор $N_t = [N_0(t), N_1(t), ..., N_m(t)]$ назовем состоянием системы в момент времени t.

Поскольку восстановление (замена) отказавших элементов связана с необходимостью прерывания процесса работы системы, то проведение частичных замен отказавших элементов или блоков чаще всего нецелесообразно. Далее будем предполагать, что если в некоторый момент принимается решение о восстановлении, то производится полная замена всех отказавших элементов и блоков идентичными в течение некоторого (случайного) времени τ. Другими словами, если в некоторый момент времени υ принимается решение о восстановлении, то к моменту времени $\upsilon + \tau$ система полностью возвращается к начальному (исправному) состоянию $N = (L, 0, ..., 0)$. На интервале времени $(\upsilon, \upsilon + \tau)$, в течение которого происходит восстановление (замена) отказавших элементов, процесс работы системы прерывается.

Множество возможных состояний (фазовое пространство) процесса N_t обозначим через

$$\Phi = \left\{ N = (N_0, N_1, ..., N_m) : \sum_{k=0}^{m} N_k = L, \ N_k = 0, 1, ..., L; \ k = 0, 1, ..., m \right\}.$$

В используемых предположениях процесс N_t является марковским с непрерывным временем и конечным множеством состояний Φ с матрицей интенсивностей переходов из состояния $N = (N_0, N_1, ..., N_m) \in \Phi$ в $N' = (N'_0, N'_1, ..., N'_m) \in \Phi$ следующего вида:

$$\lambda(N, N') = N_k (m - k) \lambda,$$

если $N'_j = N_j$ при всех $j \neq k, k+1$ и $N'_k = N_k - 1$, $N'_{k+1} = N_{k+1} + 1$, где k принимает значения 0, 1, …, $m - 1$, и

$$\lambda(N, N') = 0$$

для всех остальных $N \in \Phi$, $N' \in \Phi$.

Величина $N_k(t)$ представляет собой число блоков, находящихся в момент времени t в состоянии k. Поскольку процессы отказов $k_i(t)$ в различных блоках предполагаются независимыми, то, следовательно, случайный вектор $N_t = [N_0(t), N_1(t), ..., N_m(t)]$ имеет полиномиальное распределение. Другими словами, вероятностное распределение на множестве Φ состояний процесса N_t в момент времени t имеет вид

$$P_t(N) = P\{N_t = N\} = \frac{L!}{N_0! N_1! \cdot ... \cdot N_m!} \prod_{k=0}^{m} [p_t(k)]^{N_k}, \ N \in \Phi. \qquad (2)$$

Введем функции состояний:

$$D(N) = \sum_{k=0}^{m} kN_k; \ M(N) = \sum_{k=0}^{m} (m-k)N_k; \ R(N) = \sum_{k=1}^{m} N_k ,$$

где $D(N)$ — число отказавших элементов; $M(N)$ — число исправных элементов; $R(N)$ — число блоков, содержащих хотя бы один отказавший элемент, в состоянии $N \in \Phi$.

Траектории процесса N_t будем предполагать непрерывными справа по t. Пусть в момент времени $t = 0$ процесс N_t находится в состоянии полной исправности: $N_0 = (L, 0, ..., 0)$. Момент остановки υ процесса работы системы и начала восстановления (замены) отказавших блоков определим как момент первого достижения процессом $N_k(t)$ области «остановки» $G \in \Phi$:

$$\upsilon = \min\{t : N_t \in G\}. \qquad (3)$$

Будем предполагать, что множество G является поглощающим. На интервале $(\upsilon, \upsilon + \tau)$ работа системы прерывается и происходит восстановление (замена) отказавших блоков. Затем в момент $\upsilon + \tau$ процесс N_t снова стартует из состояния полной исправности $N_0 = (L, 0, ..., 0)$ и работа системы возобновляется до следующего попадания процесса N_t в множество остановки G, после чего снова происходит прерывание работы системы и восстановление отказавших блоков в течение времени τ и т.д. Таким образом, система функционирует последовательными независимыми циклами:

$$(\upsilon_n, \tau_n), \ n = 1, 2, ..., \qquad (4)$$

где (υ_n, τ_n) — n-я независимая реализация случайных величин (υ, τ).

Обозначим $\eta_n = \sum_{k=1}^{n} (\upsilon_k + \tau_k)$ — момент завершения n-го цикла (n-го восстановления) и $\zeta_n = \eta_{n-1} + \upsilon_n$ — момент начала восстановления на n-м цикле, где $\eta_0 = 0$; $n = 1, 2, ...$. Тогда число элементов $V(t)$, работоспособных в текущий момент времени t, записывается через определенный выше процесс $M(N_t)$ следующим образом:

$$V(t)=\begin{cases} M(N_{U(t)}) & \text{при } 0\le U(t)<\upsilon_n;\\ 0 & \text{при } \upsilon_n\le U(t)<\upsilon_n+\tau_n, \end{cases}$$

где $U(t)=t-\eta_{n-1}$ — время от момента η_{n-1} начала n-го цикла до текущего момента t.

Рассмотрим отдельный цикл функционирования процесса N_t, $t\ge 0$. Величину

$$S_t=\int_0^t M(N_u)\,du,$$

равную суммарному времени работы всех элементов на интервале $(0,t)$, будем называть суммарной наработкой элементов системы к моменту времени t, а

$$S_\upsilon=\int_0^\upsilon M(N_t)\,dt$$

— суммарной наработкой всех элементов от начального момента цикла $t=0$ до момента восстановления υ.

Среднее число работающих элементов в системе на интервале $(0,\infty)$

$$\overline{V}=\lim_{T\to\infty}\frac{1}{T}E\int_0^T V(t)dt=\frac{ES_\upsilon}{E\upsilon+E\tau}, \tag{5}$$

где E — символ математического ожидания.

Выражение (5) позволяет свести вычисление характеристики $\overline{V}$ на бесконечном интервале времени $(0,\infty)$ к характеристикам $E\tau$, $E\upsilon$, ES_υ процесса N_t на одном цикле.

Пусть $g(N)$ — штраф (затраты) на восстановление, если решение о восстановлении принимается в момент, когда система находится в состоянии $N\in\Phi$. Будем предполагать, что штрафная функция имеет вид

$$g(N)=\sum_{k=1}^m C_k N_k, \tag{6}$$

где C_k — стоимость восстановления (замены) одного блока, находящегося в состоянии k, т. е. имеющего в своем составе k отказавших элементов.

В частности, если $C_k = b$ (где b — стоимость замены одного блока), то штрафная функция (6) имеет вид

$$g(N) = b\sum_{k=1}^{m} N_k = b\,R(N),$$

что соответствует ситуации замен, когда при восстановлении полностью заменяются все блоки, содержащие хотя бы один отказавший элемент. Если $C_k = \beta k$ (где β — стоимость замены одного элемента), то в этом случае

$$g(N) = \beta\sum_{k=1}^{m} k\,N_k = \beta\,D(N),$$

что соответствует ситуации замен, когда при восстановлении заменяются только все отказавшие элементы (без замен блоков, содержащих отказавшие элементы).

Аналогичным образом штрафная функция (6) описывает и более общие правила восстановления, например, когда при восстановлении проводится замена всех отказавших элементов, а также тех блоков, в которых число отказавших элементов превосходит некоторый критический уровень k^*. В этом случае $C_k = \beta k$ при $k < k^*$ и $C_k = b$ при $k \geq k^*$.

При выбранной штрафной функции $g(N)$

$$\overline{g} = \lim_{n\to\infty}\frac{g(N_{\upsilon_1}) + g(N_{\upsilon_n}) + \ldots + g(N_{\upsilon_n})}{(\upsilon_1+\tau_1)+(\upsilon_2+\tau_2)+\ldots+(\upsilon_n+\tau_n)} = \frac{Eg(N_\upsilon)}{E\upsilon + E\tau} \qquad (7)$$

характеризует средние затраты на замену отказавшего оборудования в единицу времени на интервале $(0,\infty)$. Выражение (7), как и приведенная выше формула (5), позволяет свести вычисление характеристики $\overline{g}$ качества функционирования системы на бесконечном интервале времени $(0,\infty)$ к вычислению характеристик $E\upsilon$, $E\tau$, $Eg(N_\upsilon)$ процесса N_t на одном цикле $(0,\upsilon)$.

Таким образом, правило восстановления отказавшего оборудования задается марковским моментом остановки (прерывания) υ процесса работы системы и начала восстановления на каждом очередном цикле. Момент υ, в свою очередь, задается областью «остановки» G или областью «продолжения работы» $H = \Phi\setminus G.$ На каждом очередном цикле работа системы начинается из состояния полной исправности N_0 и продолжается до тех пор, пока процесс N_t находится в области H, прерываясь на время τ при первом достижении процессом N_t области «остановки» G. Далее задача сводится к вычислению по формулам (5) и (7) указанных выше основных показателей качества работы системы в зависимости от той или иной выбранной области «остановки» (начала восстановления) G в фазовом пространстве состояний системы Φ. Зависимость этих показателей от выбора областей G и $H = \Phi \setminus G$ обозначим как $\overline{V} = \overline{V}_H$, $\overline{g} = \overline{g}_H$. Определение оптимального правила начала восстановления далее заключается в максимизации показателя эффективности

$$\max \overline{V}_H, \; H \subset \Phi, \tag{8}$$

при ограничении на связанные с восстановлением средние затраты

$$\overline{g}_H \le g^*, \; H \subset \Phi, \tag{9}$$

где g^* — заданный критический уровень допустимых затрат.

Вычисление основных показателей. Приведенная ниже теорема позволяет достаточно просто вычислять указанные выше основные показатели $E\upsilon$, ES_υ, $Eg(N_\upsilon)$ в выражениях (5) и (7), по крайней мере для случая, когда область «остановки» G является поглощающей. Пусть $f(N) \ge 0$ — некоторая функция состояния $N \in \Phi$. Выражение

$$E\int_0^{\upsilon} f(N_t)dt \tag{10}$$

назовем обобщенной средней суммарной наработкой всех элементов системы к моменту начала восстановления υ. Нетрудно увидеть, что характеристики $E\upsilon$ и ES_υ являются частными случаями выражения (10) соответственно при $f(N) \equiv 1$ и $f(N) = M(N)$.

Теорема. Если область «остановки» G в выражении (5) является поглощающей, то

$$E\int_0^{\upsilon} f(N_t)dt = \sum_{N \in H} f(N) \int_0^{\infty} P_t(N)dt; \tag{11}$$

$$Eg(N_\upsilon) = \sum_{K \in G} g(K) \sum_{N \in H} \lambda(N, K) \int_0^{\infty} P_t(N)dt, \tag{12}$$

где $H = \overline{G} = \Phi \setminus G$ — область «продолжения работы» системы.

Доказательство. С учетом конечности множества Φ имеем

$$E\int_0^{\upsilon} f(N_t)dt = E\int_0^{\infty} I(\upsilon > t) f(N_t)dt =$$

$$= E\int_0^{\infty} I(\upsilon > t) \sum_{N\in\Phi} f(N) I(N_t = N)\, dt =$$

$$= \sum_{N\in\Phi} f(N) E\int_0^{\infty} I(\upsilon > t)\, I(N_t = N)\, dt, \qquad (13)$$

где $I(A)$ — индикатор события A.

Согласно определению момента остановки υ (см. выражение (3)), при любом $t \ge 0$ справедливо соотношение включения

$$\{\upsilon > t\} \subset \{N_t \in H\}.$$

Кроме того, если область G поглощающая, то для этих событий справедливо и обратное соотношение

$$\{N_t \in H\} \subset \{\upsilon > t\}$$

при любом $t \ge 0$. Тем самым, если область G поглощающая, то при любом $t \ge 0$ справедливо выражение

$$I(\upsilon > t) = I(N_t \in H). \qquad (14)$$

Тогда из выражений (13) и (14) следует

$$E\int_0^{\upsilon} f(N_t) dt = \sum_{N\in\Phi} f(N) E\int_0^{\infty} I(N_t \in H) I(N_t = N)\, dt =$$

$$= \sum_{N\in\Phi} f(N)\, E\int_0^{\infty} I(N_t = N)\, dt = \sum_{N\in\Phi} f(N) \int_0^{\infty} P_t(N)\, dt,$$

т. е. равенство (11) доказано. Заметим, что в частном случае (при $f(N) \equiv 1$) равенство (11) с учетом (14) имеет вид

$$E\upsilon = \sum_{N\in H} \int_0^{\infty} P_t(N) dt = \int_0^{\infty} P_t\{N_t \in H\}\, dt = \int_0^{\infty} P_t\{\upsilon > t\}\, dt,$$

который совпадает с известным выражением для математического ожидания момента υ первого достижения поглощающего множества состояний $G = \Phi \setminus H$.

Для доказательства равенства (12) рассмотрим разбиение временной оси $t \ge 0$ дискретными моментами $t_j = jh$, $j = 0, 1, 2, \ldots$, где h — шаг разбиения. Поскольку область G поглощающая, то в соответствии с (13) при любом $t_j \ge 0$ справедливо следующее равенство между событиями:

$$\{\upsilon > t_j\} = \{N(t_j) \in H\}. \qquad (15)$$

Пусть $K \in G$. Тогда событие $\{N_\upsilon = K\}$ может быть представлено в виде

$$\{N_\upsilon = K\} = \sum_{j=0}^{\infty} A_j(K), \tag{16}$$

где $A_j(K) = \{t_j < \upsilon \le t_j + h\} \cap \{N_\upsilon = K\}$.

С учетом равенства (15) справедливо соотношение

$$\{t_j < \upsilon \le t_j + h\} \subset \{\upsilon > t_j\} = \{N(t_j) \in H\}.$$

Отсюда получаем, что событие $A_j(K)$ далее может быть представлено в виде

$$A_j(K) = \{N(t_j) \in H\} \cap \{t_j < \upsilon \le t_j + h\} \le \{N_\vartheta = K\}, \tag{17}$$

т. е. событие $A_j(K)$ состоит в том, что на интервале времени $(t_j, t_j + h)$ происходит переход процесса из некоторого состояния $N \in H$ в состояние $K \in G$. С учетом равенства (17), а также соотношения

$$\{N(t_j) \in H\} = \sum_{N \in H} \{N(t_j) = N\}$$

событие (16) может быть представлено в виде

$$\{N_\upsilon = K\} = \sum_{j=0}^{\infty} \sum_{N \in H} B_j(N, K), \tag{18}$$

где $B_j(N, K) = \{N(t_j) = N\} \cap \{t_j < \upsilon \le t_j + h\} \le \{N_\upsilon = K\}$ — событие, состоящее в том, что в момент времени t_j процесс $N(t)$ находится в состоянии $N \in H$ и на интервале $(t_j, t_j + h)$ происходит его переход в состояние $K \in G$.

Пусть $T > 0$ — некоторый фиксированный момент времени. Выберем шаг h разбиения временной оси так, что $n = T/h$ — целое число. События $B_j(N, K)$ являются непересекающимися при всех различных $j = 0, 1, \ldots$, $N \in H$, $K \in G$. Из выражения (18) следует

$$P\{N_\upsilon = K\} = \sum_{j=0}^{n-1} \sum_{N \in H} P\{B_j(N, K)\} + \sum_{j=n}^{n-1} \sum_{N \in H} P\{B_j(N, K)\}. \tag{19}$$

Вторая сумма в выражении (19) заведомо не превосходит $P\{\upsilon > T\}$. Имеем

$$P\{B_j(N,K)\} = P\{N(t_j = N)\}[\lambda(N,K)h + \xi],$$

где $|\xi| < C_1 h^2$; C_1 — константа, не зависящая от $j = 0,1,...$; $N \in H$; $K \in G$.

Отсюда с учетом выражения (19)

$$P\{N_\upsilon = K\} = \sum_{N\in H}\sum_{j=0}^{n-1} P\{N(t_j) = N\}\lambda(N,K)h + \delta_1 + \delta_2. \qquad (20)$$

Здесь

$$\delta_1 \le C_1 C_2 n h^2 = C_1 C_2 T h;\ \delta_2 \le P\{\upsilon \ge T\} \le e^{-\lambda T}; \qquad (21)$$

$C_2 < \infty$ — общее число всех возможных состояний процесса $N(t) \in \Phi$.

Поскольку $h > 0$, $T > 0$ — произвольные константы (причем $n = T/h$ — целое число), то из выражений (20), (21), учитывая также непрерывность функции $P_t(N) = P\{N(t) = N\}$ по t, следует равенство (12). Теорема доказана.

Таким образом, полученные соотношения позволяют вычислять среднее число работоспособных элементов в системе и средние затраты на восстановление в системе в единицу времени в стационарном режиме на интервале $(0,\infty)$. На основе этих выражений можно строить соответствующие численные алгоритмы для решения задач оптимизации (см. формулы (8) , (9)). Отметим также, что существенный интерес представляет дальнейшее обобщение приведенных результатов на более общие модели, учитывающие такие факторы, как зависимость элементов, возможность кратных отказов элементов внутри отдельных блоков (подсистем), возможность «блочных» отказов в системе и др.

СПИСОК ЛИТЕРАТУРЫ

1. Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. – М.: Наука, 1965. – 524 с.
2. Gnedenko B.V., Pavlov I.V., Ushakov I.A. Statistical reliability engineering. – N.Y.: John Wiley & Sons, 1999. – 514 p.
3. Ронжин А.Ф., Суриков В.Н. О времени полного перебора // Обозрение прикладной и промышленной математики. – 2007. – Т. 14. – № 3. – С. 506–508.
4. Коновалов М.Г., Малашенко Ю.Е., Назарова И.А. Модели и методы управления заданиями в системах распределенных вычислительных ресурсов – М.: Вычислительный центр РАН, «Сообщения по прикладной математике», 2009. – 110 с.

5. Лёвин П.А., Павлов И.В. Оценка надежности систем радиоэлектроники в переменном режиме функционирования // Успехи современной радиоэлектроники. – 2009. – № 3. – С. 73–79.
6. Лёвин П.А, Павлов И.В. Оценка надежности системы с нагруженным резервированием по результатам испытаний ее элементов //Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. – 2011. – № 3. – С. 59–70.

Статья поступила в редакцию 28.09.2012