Distribution of operational times in fault-tolerant systems modeled by semi-Markov reward processes

Gerardo Rubino, Bruno Sericola

INRIA

UNITE DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 996

Programme 3

DISTRIBUTION OF OPERATIONAL TIMES IN FAULT-TOLERANT SYSTEMS MODELED BY SEMI-MARKOV REWARD PROCESSES

Gerardo RUBINO
Bruno SERICOLA

Mars 1989

RR-0996

IRISA INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CÉDEX
FRANCE
Téléphone : 99 36 20 00
Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

Distribution of operational times in fault-tolerant systems modeled by semi-Markov reward processes

Gerardo Rubino
Bruno Sericola

Publication Interne n° 456 - Février 1989 - 10 pages

Abstract :

We consider fault-tolerant systems with operational periods in which the system works and repair periods in which the system performs a reconfiguration from a recovering point. We are interested in obtaining informations about the successive operational periods before a fatal breakdown. In order to be able to deal with arbitrary probability distributions for the holding times in each state of the system, we assume that it is modeled by a semi-Markov process. In this paper we give the distribution of the n^{th} operational period. It often happens that the user wants to make a distinction among the operational states in order to get more detailed information. We allow then the use of rewards on the model states, that is, the results will be presented for semi-Markov reward processes. An example will illustrate this work.

Distribution des temps opérationnels dans les systèmes tolérant les pannes modélisés par des processus semi-markoviens à taux de récompense

Résumé :

On considère des systèmes tolérant les pannes avec des périodes opérationnelles dans lesquelles le système fonctionne correctement et des périodes de réparation dans lesquelles le système tente une reconfiguration à partir d'un point de reprise. On cherche à obtenir des informations sur les périodes opérationnelles successives précédent une erreur fatale. On suppose le système modélisé par un processus semi-markovien, de manière à pouvoir traiter des distributions de probabilité quelconques pour les temps de séjour dans chacun de ses états. Dans ce rapport, on donne la distribution de la $n^{ième}$ période opérationnelle. Il apparaît souvent que l'utilisateur désire différencier les états opérationnels pour obtenir des informations plus détaillées. On permet alors l'utilisation de taux de récompense sur les états du modèle, les résultats sont donc présentés pour des processus semi-markoviens à taux de récompense. Un exemple illustre ce travail.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(L. A. 227)
UNIVERSITÉ DE RENNES 1 I. N. S. A. DE RENNES

INSTITUT NATIONAL DE RECHERCHE
EN INFORMATIQUE ET EN AUTOMATIQUE
(LABORATOIRE DE RENNES)

1 Model description

We consider a fault-tolerant system which is able to recover from breakdowns. At any instant t, there are three possibilities. The system can be operational, that is, performing the tasks assigned to it (perhaps in a degradated way). It can be attempting to reconfigure itself after a failure without doing any useful work. Finally, it can be dead, after some fatal breakdown.

The system is modeled by a right continuous homogeneous semi-Markov process denoted by $X = \{X_t, t \geq 0\}$. The semi-Markov assumption, instead of considering Markov processes, allows us to use general probability distributions in the individual states. The finite state space, denoted by E, is assumed to be composed by transient states and recurent states. Since we are interested in transient measures, all the recurrent classes can be lumped into only one absorbing state. That is, we suppose that $E = \{1, ..., N, a\}$ where $1, \ldots, N$ are transient and a is absorbing. Let T_k be the time of the k^{th} transition ($T_0 = 0$) and define $V_k = T_{k+1} - T_k$, the sojourn time in the $(k+1)^{th}$ visited state ($k \in \mathbb{N}$). Denote by $X(k) = X_{T_k}$ the state reached after the k^{th} transition. The process X is defined by its kernel Q_t and its initial probability distribution α. The transition probability matrix of the embedded Markov chain $\{X(k), k \in \mathbb{N}\}$ is denoted by P. For every $i, j \in E$, we have

$$Q_t(i,j) \stackrel{\text{def}}{=} \mathbb{P}(X(k+1) = j, V_k \leq t | X(k) = i)$$

$$P(i,j) \stackrel{\text{def}}{=} Q_\infty(i,j) = \mathbb{P}(X(k+1) = j | X(k) = i)$$

We assume for instance that $Q_t(a,a) = \delta(t-1)$, where

$$\delta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Note also that $Q_t(a,j) = 0$ for every $j \neq a$.

A non-negative real-valued reward rate r_i is associated with each state $i \in E$. This allows the modeler to differenciate the states of B and, in this way, to capture not only failure and repair impact but also, for instance, different performance levels.

We denote by B the subset of E containing the operational states and by B' the subset of E containing the others transient states, so we have: $E = B \cup B' \cup \{a\}$. With the operational states we associate strictly positive reward rates. We will see that the reward rates associated with the others states has no influence on the results.

The distribution of the n^{th} sojourn time in a subset of states for an irreducible and homogeneous Markov process can be found in [1]. Here the process is neither irreducible nor Markov and has reward rates. However, the two results are closely related. For $n \geq 1$, let $S_{i,B,n}$ denote the total time that X spends in state $i \in B$ during its n^{th} visit to B. In the next section, we analyze the random variable $S_{B,n} \stackrel{\text{def}}{=} \sum_{i \in B} r_i S_{i,B,n}$. The distribution of $S_B \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} S_{B,n}$ which represents the total accumulated reward until absorption can be found in [2]. Section 3 presents a short application.

In the sequel, we will denote by 1^T the column vector with all its coordinates equal to 1 (we will always use row vectors and $(.)^T$ denotes the transpose operator) and by I the identity matrix, their dimensions being defined by the context.

2 Distribution of $S_{B,n}$

Let us consider another semi-Markov process without rewards, denoted by $Y = \{Y_t, t \geq 0\}$, defined on the state space $B \cup \{b\}$ where the state b is absorbing. The initial probability distribution of Y is $(\alpha_B, 1 - \alpha_B 1^T)$ where α_B denotes the subvector of α corresponding to the states of B. The kernel $\hat{Q}_t$ of Y is defined as follows.

$$\begin{array}{lcll} \hat{Q}_t(i,j) & \stackrel{\text{def}}{=} & Q_{t/r_i}(i,j) & \text{for } i,j \in B \\ \hat{Q}_t(i,b) & \stackrel{\text{def}}{=} & Q_{t/r_i}(i,a) + \sum_{j \in B'} Q_{t/r_i}(i,j) & \text{for } i \in B \\ \hat{Q}_t(b,j) & \stackrel{\text{def}}{=} & Q_t(a,j) & \text{for } j \in B \\ \hat{Q}_t(b,b) & \stackrel{\text{def}}{=} & Q_t(a,a) & \end{array}$$

We are now able to derive the conditional distribution of $S_{B,1}$ given that the initial state is in the subset B.

Lemma 2.1 *For every $i \in B$, $\mathbb{P}(S_{B,1} \leq t / X_0 = i) = \mathbb{P}(Y_t = b / Y_0 = i)$.*

Proof. For every $i \in B$,

$$\begin{aligned} \mathbb{P}(S_{B,1} \leq t / X_0 = i) &= \sum_{j \in B} \int_0^{t/r_i} \mathbb{P}(S_{B,1} \leq t - r_i s / X_0 = j) dQ_s(i,j) \\ &\quad + \sum_{j \in B'} Q_{t/r_i}(i,j) + Q_{t/r_i}(i,a) \\ &= \sum_{j \in B} \int_0^t \mathbb{P}(S_{B,1} \leq t - s / X_0 = j) d\hat{Q}_s(i,j) + \hat{Q}_t(i,b). \end{aligned}$$

On the other hand, we have

$$\mathbb{P}(Y_t = b / Y_0 = i) = \sum_{j \in B} \int_0^t \mathbb{P}(Y_{t-s} = b / Y_0 = j) d\hat{Q}_s(i,j) + \hat{Q}_t(i,b).$$

So, these two quantities are solutions to the same integral equation. Since this equation has an unique solution (see [3]), the enounced result follows. □

Let us decompose the matrix P and the initial probability vector α with respect to the partition $E = B \cup B' \cup \{a\}$ as follows.

$$P = \begin{pmatrix} P_B & P_{BB'} & P_{Ba} \\ P_{B'B} & P_{B'} & P_{B'a} \\ 0 & 0 & 1 \end{pmatrix}, \quad \alpha = (\alpha_B \ \ \alpha_{B'} \ \ \alpha_a).$$

We denote by H the $B' \times B$ matrix defined by $H \stackrel{\text{def}}{=} (I - P_{B'})^{-1} P_{B'B}$. We define the vectors

$$\begin{aligned} u_B(n,t) &\stackrel{\text{def}}{=} (\mathbb{P}(S_{B,n} \leq t/X_0 = i), i \in B), \\ u_{B'}(n,t) &\stackrel{\text{def}}{=} (\mathbb{P}(S_{B,n} \leq t/X_0 = i), i \in B'). \end{aligned}$$

With this notation, we prove the following lemma.

Lemma 2.2 *For all $n \geq 1$, $u_{B'}^T(n,t) = 1^T - H\left(1^T - u_B^T(n,t)\right)$.*

Proof. For every $i \in B'$,

$$\begin{aligned} \mathbb{P}(S_{B,n} \leq t/X_0 = i) &= \sum_{j \in B} P(i,j)\mathbb{P}(S_{B,n} \leq t/X_0 = j) \\ &\quad + \sum_{j \in B'} P(i,j)\mathbb{P}(S_{B,n} \leq t/X_0 = j) + P(i,a) \end{aligned}$$

In matrix notation, we obtain

$$u_{B'}^T(n,t) = P_{B'B} u_B^T(n,t) + P_{B'} u_{B'}^T(n,t) + P_{B'a}$$

which can easily be written

$$u_{B'}^T(n,t) = 1^T - H\left(1^T - u_B^T(n,t)\right)$$

since $P_{B'B}1^T + P_{B'}1^T + P_{B'a} = 1^T$. □

Let us denote by G the $B \times B$ matrix $G \stackrel{\text{def}}{=} (I - P_B)^{-1} P_{BB'} H$, and by v_1 the vector $v_1 \stackrel{\text{def}}{=} \alpha_B + \alpha_{B'} H$. We get the following expression of the distribution of $S_{B,n}$.

Theorem 2.3 *For all $n \geq 1$, $\mathbb{P}(S_{B,n} \leq t) = 1 - v_1 G^{n-1}\left(1^T - u_B^T(1,t)\right)$.*

Proof. Let $i \in B$ and $n \geq 2$.

$$\begin{aligned} \mathbb{P}(S_{B,n} \leq t/X_0 = i) &= \sum_{j \in B} P(i,j)\mathbb{P}(S_{B,n} \leq t/X_0 = j) \\ &\quad + \sum_{k \in B'} P(i,k)\mathbb{P}(S_{B,n-1} \leq t/X_0 = k) + P(i,a) \end{aligned}$$

which gives in matrix notation, using Lemma 2.2,

$$\begin{aligned} u_B^T(n,t) &= P_B u_B^T(n,t) + P_{BB'} u_{B'}^T(n-1,t) + P_{Ba} \\ &= 1^T - G1^T + G u_B^T(n-1,t). \end{aligned}$$

It follows that, for all $n \geq 1$,

$$\begin{aligned} 1^T - u_B^T(n,t) &= G\left(1^T - u_B^T(n-1,t)\right) \\ &= G^{n-1}\left(1^T - u_B^T(1,t)\right). \end{aligned}$$

Finally, we obtain

$$\begin{aligned}
\mathbb{P}(S_{B,n} \leq t) &= \alpha_B u_B^T(n,t) + \alpha_{B'} u_{B'}^T(n,t) + \alpha_a \\
&= 1 - \alpha_B(1^T - u_B^T(n,t)) - \alpha_{B'}(1^T - u_{B'}^T(n,t)) \\
&= 1 - v_1(1^T - u_B^T(n,t)) \quad \text{using Lemma 2.2} \\
&= 1 - v_1 G^{n-1}(1^T - u_B^T(1,t)).
\end{aligned}$$

Consider the sequence of states in which the successive visits of X to B begin. If we add to the end of each sequence the absorbing state a for each trajectory in which X is absorbed (this happens with probability 1), a homogeneous discrete time Markov chain is defined with state space $B \cup \{a\}$. It can be shown that the submatrix of the transition probability matrix of this chain obtained by deleting the row and column corresponding to the absorbing state a, is G. In the same way, $(v_1\ ,\ 1 - v_1 1^T)$ is its starting probability distribution. The reader is referred to [1] for a proof in a similar context (irreducible Markov processes without rewards). □

Note that the vector $u_B^T(1,t)$ is given by Lemma 2.1. A numerical inversion of Laplace transform can be used to compute it (see for example [4]).

3 Application

As an application of the previous results, we consider a fault-tolerant architecture in which three identical processors work in parallel running the same code. The three processors receive the same input and take synchroneously a binary decision. A voter device controls the three outputs and in case of disagreement between them, it chooses the value which is present twice. There can be software or hardware faults on the processors and the voter is assumed here perfectly reliable. The system starts with the three processors performing correctly (state 1 in Figure 1). Due to design faults in the software, the system can be shut down and a recovery procedure is started (state 4). We assume that the software faults occur with rate λ_s and that the repairing time is constant and lasts K. The probability of a software fault recovery is denoted by d (the *software fault coverage* parameter).

With respect to the hardware faults, the three processor behaviours are assumed to be independent. The hardware faults occur at rate λ_h for any processing unit. When such an error occurs, a procedure try to put the system back in operation. Its success probability is constant c (the *hardware fault coverage* parameter). This procedure is assumed to be executed instanstaneously. So, with probability c, the system continue to perform (in a degradated mode) with only one processor, the other non failed unit staying in passive redundancy (state 2). The voter is no more useless in this situation and, for simplicity, the system is assumed to be unable to recover from software errors when there is only one active processor. The redundant unit can not fail and when there is a new hardware breakdown, a

second procedure is started to try to reconfigure again. If it is successful, the system uses the remaining operational processor (state 3) until the last (fatal) hardware fault (state 5).

This model leads to the five-states semi-Markov process described in Figure 1, where the holding time in state 4 is constant and equal to K. The other holding times are exponentially distributed (exepting, of course, for state 5 which corresponds to the crash of the system). The set of operational states is then $B = \{1, 2, 3\}$.

Assume that an active unit gives the wrong answer with probability p. When three processors are working in parallel (and independently), the probability of obtaining a wrong output is $p^3 + 3p^2(1-p)$. Furthermore, it is assumed that the workload of the system is high, that is, the system is required to give a large number of outputs per unit of time. So, we consider the following reward rates for the states of B: $r_1 = 1 - 3p^2 + 2p^3$, $r_2 = r_3 = r = 1 - p$. With these assumptions, the random variables $S_{B,n}$ can be viewed as the total time while the system gives the right answer during the n^{th} operational period.

$$Q_t(1,2) = \frac{3\lambda_h c}{3\lambda_h+\lambda_s}(1 - e^{-(3\lambda_h+\lambda_s)t})$$

$$Q_t(1,4) = \frac{\lambda_s}{3\lambda_h+\lambda_s}(1 - e^{-(3\lambda_h+\lambda_s)t})$$

$$Q_t(1,5) = \frac{3\lambda_h(1-c)}{3\lambda_h+\lambda_s}(1 - e^{-(3\lambda_h+\lambda_s)t})$$

$$Q_t(2,3) = \frac{\lambda_h c}{\lambda_h+\lambda_s}(1 - e^{-(\lambda_h+\lambda_s)t})$$

$$Q_t(2,5) = \frac{\lambda_h(1-c)+\lambda_s}{\lambda_h+\lambda_s}(1 - e^{-(\lambda_h+\lambda_s)t})$$

$$Q_t(3,5) = 1 - e^{-(\lambda_h+\lambda_s)t}$$

$$Q_t(4,1) = d\delta(t - K)$$

$$Q_t(4,5) = (1-d)\delta(t - K)$$

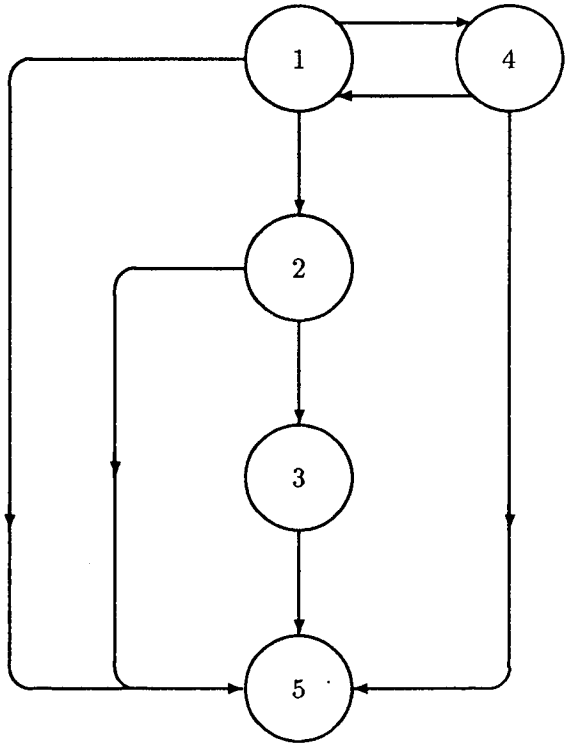

Figure 1: A three-processors fault-tolerant architecture

The computations can be easily made by hand since the holding times in the states of B are exponentially distributed. The distribution of the n^{th} accumulated reward $S_{B,n}$ is given by the following

expression.

$$\mathbb{P}(S_{B,n} \leq t) = 1 - \left(\frac{d\lambda_s}{3\lambda_h + \lambda_s}\right)^{n-1} \left[(1+U)\, e^{-\frac{3\lambda_h + \lambda_s}{r_1} t} - \left(\frac{3\lambda_h^2 c^2}{L} t + U\right) e^{-\frac{\lambda_h + \lambda_s}{r} t}\right]$$

where $L = \lambda_h(r_1 - 3r) + \lambda_s(r_1 - r)$ and $U = \dfrac{3\lambda_h cr}{L} + \dfrac{3\lambda_h^2 c^2 r_1 r}{L^2}$.

If we compute the corresponding distribution for the set B', reduced here to the state 4, we obtain

$$\mathbb{P}(S_{B',n} \leq t) = 1 - \frac{\lambda_s}{3\lambda_h + \lambda_s} \left(\frac{d\lambda_s}{3\lambda_h + \lambda_s}\right)^{n-1} \delta(K - t).$$

Observe that if $p = 0$ (i.e. $r_1 = r = 1$), $S_{B,n}$ represents the time spent by X in B during its n^{th} visit to this set. Given that the system starts in a state of B, the expectation of the time spent until the end of the M^{th} visit to B is then

$$\left(\sum_{n=1}^{M} \mathbb{E}(S_{B,n}) + \sum_{n=1}^{M-1} \mathbb{E}(S_{B',n})\right)\Bigg|_{p=0}$$

As $\sum_{n=1}^{M} \mathbb{E}(S_{B,n})$ represents the average total time while the system works and gives the right answer during the M first operational periods, a quantity of interest is the ratio

$$\rho(M) = \frac{\sum_{n=1}^{M} \mathbb{E}(S_{B,n})}{\left(\sum_{n=1}^{M} \mathbb{E}(S_{B,n}) + \sum_{n=1}^{M-1} \mathbb{E}(S_{B',n})\right)\Big|_{p=0}} \quad \text{and} \quad \rho(1) \stackrel{\text{def}}{=} 1$$

which represents the expected rate of right answers during the M first operational periods. In this case, we obtain for every $M \geq 2$,

$$\rho(M) = \frac{F_p}{F_0 + \frac{1-q^{M-1}}{1-q^M} \frac{q}{d} K}$$

where $F_p = \dfrac{1 - 3p^2 + 2p^3}{3\lambda_h + \lambda_s} + (1-p)\left(\dfrac{3\lambda_h c}{(\lambda_h + \lambda_s)(3\lambda_h + \lambda_s)} + \dfrac{3\lambda_h^2 c^2}{(\lambda_h + \lambda_s)^2(3\lambda_h + \lambda_s)}\right)$ and $q = \dfrac{\lambda_s}{3\lambda_h + \lambda_s}$.

See that $F_0 \geq F_p$ and that for any fixed value of K, we have $\dfrac{F_p}{F_0 + \frac{q}{d}K} < \rho(M) \leq 1$.

For a given level β with $0 < \beta < 1$ (and $\beta \approx 1$), we can evaluate, for instance, the maximal value of the parameter K (the execution time of the software recovery procedure) such that $\rho(M) \geq \beta$. Observe that if the number M of operational periods is fixed and $M \geq 2$, the best theorically possible value for the ratio $\rho(M)$ is F_p/F_0. If $\beta < F_p/F_0$ we have

$$\rho(M) \geq \beta \text{ for every } M \geq 2 \Longleftrightarrow 0 < K \leq \frac{F_p - \beta F_0}{\beta} \frac{d}{q}$$

References

[1] G. Rubino and B. Sericola. *Sojourn Time in Finite Markov Processes.* Technical Report 818, I.N.R.I.A., Campus de Beaulieu, 35042 Rennes Cedex, France, March 1988. To appear in J. of Appl. Prob., December 1989.

[2] G. Ciardo, R. Marie, B. Sericola, and K. Trivedi. *Performability analysis using semi-Markov process.* Technical Report CS-1988-9, Duke University, Comp. Sc. Dept., 1988. To appear in IEEE Trans. on Comp.

[3] E. Cinlar. *Introduction to stochastic Processes.* Prentice Hall, New-Jersey, 1975.

[4] D. L. Jagerman. An inversion technique for the Laplace transforms. *Bell System Technical Journal*, 61:1995–2002, September 1982.

LISTE DES DERNIERES PUBLICATIONS INTERNES

PI 451 **LANCER DE RAYON SUR DES ARCHITECTURES PARALLELES : UNE ETUDE DE PERFORMANCE**
Thierry PRIOL, Kadi BOUATOUCH
20 Pages, Janvier 1989.

PI 452 **LANCER DE RAYON : APPROCHES PARALLELES**
Didier BADOUEL, François BODIN, Thierry PRIOL
16 Pages, Janvier 1989.

PI 453 **UN COMPILATEUR ESTELLE MULTI-PROCESSEURS POUR L'EXPERIMENTATION D'ALGORITHMES DISTRIBUES SUR MACHINES PARALLELES**
Jean-Marc JEZEQUEL, Claude JARD
54 Pages, Janvier 1989.

PI 454 **REALISATION ET CALIBRATION D'UN SYSTEME EXPERIMENTAL DE VISION COMPOSE D'UNE CAMERA MOBILE EMBARQUEE SUR UN ROBOT-MANIPULATEUR**
François CHAUMETTE, Patrick RIVES
36 Pages, Février 1989.

PI 455 **ARCHITECTURE SYSTOLIQUE POUR LA CORRECTION AUTOMATIQUE DE LIBELLE D'ADRESSE**
Dominique LAVENIER, Jean-Luc SCHARBARG, Patrice FRISON
22 Pages, Février 1989.

PI 456 **DISTRIBUTION OF OPERATIONAL TIMES IN FAULT-TOLERANT SYSTEMS MODELED BY SEMI-MARKOV REWARD PROCESSES**
Gerardo RUBINO, Bruno SERICOLA
10 Pages, Février 1989.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique