Journal of the Operations Research
Society of Japan
Vol. 42, No. 4, December 1999

CUMULATIVE DAMAGE MODEL WITH TWO KINDS OF SHOCKS AND ITS APPLICATION TO THE BACKUP POLICY

Cunhua Qian
Aichi Institute of Technology

Syouji Nakamura
The Bank of Nagoya, Ltd.

Toshio Nakagawa
Aichi Institute of Technology

(Received February 8, 1999; Revised May 14, 1999)

Abstract This paper considers an extended cumulative damage model with two kinds of shocks: One is failure shock at which a system fails and the other is damage shock at which it suffers only damage. Shocks occur at a nonhomogeneous Poisson process. A system fails when a failure shock occurs or the total damage has exceeded a threshold level K. A system is also replaced before failure at scheduled time T. Reliability measures of this model are derived, using the theory of cumulative processes. Further, this is applied to the backup of files in a database system. Optimal replacement times which minimize the expected cost are discussed and numerically computed for several cases.

1. Introduction

Shocks occur at a random point and each shock causes an amount of damage to a system. These damages accumulate additively and a system fails when the total amount of damage has exceeded a threshold level K. Such a stochastic model generates a cumulative process [3]. Some aspects of damage models from reliability viewpoints were discussed by Esary, Marshall and Proschan [5].

It is of great interest that a system is replaced before failure as preventive maintenance. The replacement policies where a system is replaced before failure at time T [17], at shock N [10], or at damage Z [6, 9] were considered. Nakagawa and Kijima [11] applied the periodic replacement with minimal repair [1] at failure to a cumulative damage model and obtained optimal values T^*, N^* and Z^* which minimize the expected cost.

This paper considers a cumulative damage model with two kinds of shocks described in Figure 1 and Figure 2: A system suffers two kinds of shocks which occur at a nonhomogeneous Poisson process. We call that one is *failure shock* at which a system fails and the other is *damage shock* at which it suffers only damage. These damages accumulate additively and a system also fails when the total damage has exceeded a threshold level K. A system is replaced at failure (see Figure 1). However, to lessen a replacement cost after failure, a system is also replaced before failure at scheduled time T as preventive maintenance (see Figure 2).

In this paper, we apply this cumulative damage model to the backup of files in a database system [16]. We suggest a stochastic backup model of files, by putting *damage* by *dumped files*, *damage shock* by *update* and *failure shock* by *database failure*. We obtain the expected cost rate $C(T)$, and discuss an optimal full backup time T^* which minimizes $C(T)$, when a database is updated at a Poisson process. It is shown that an optimal T^* is determined by a unique solution of an equation.

Further, Yeh [19] has considered the replacement model where the failure times of a system after repairs form a geometric process. We obtain the expected cost of the backup

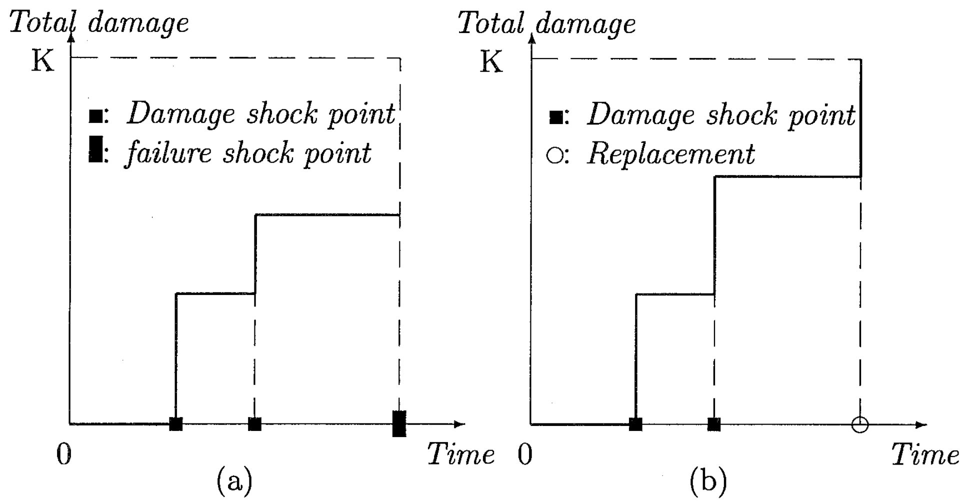

Figure 1: Replacement of cumulative damage model

model where each dumped file due to updates has a different distribution, and explain why these files increase in a geometric ratio. Finally, numerical examples are given for the backup of files in a database system when the dumped files at each update have identical exponential and different exponential distributions.

2. Damage Model

Suppose that shocks occur at a nonhomogeneous Poisson process with an intensity function $\gamma(t)$ and a mean-value function $\Gamma(t)$, *i.e.*, $\Gamma(t) \equiv \int_0^t \gamma(u)du$. Further, it is assumed that the probability that the damage shock occurs is p $(0 < p \leq 1)$ and the probability that failure shock occurs is $1-p$. It is noted that failure shocks occur at a nonhomogeneous Poisson process with an intensity function $(1-p)\gamma(t)$, and damage shocks occur at a nonhomogeneous Poisson process with an intensity function $\lambda(t) \equiv p\gamma(t)$ and a mean-value function $R(t) \equiv p\Gamma(t)$ [12]. Then, the probability that the j-th damage shock occurs exactly during $(0, t]$ is

$$H_j(t) \equiv \frac{[R(t)]^j}{j!}e^{-R(t)} \quad (j = 0, 1, 2, \cdots), \tag{2.1}$$

where $R(0) \equiv 0$.

Further, an amount Y_j of damage due to the j-th damage shock has a probability distribution $G_j(x) \equiv P_r\{Y_j \leq x\}$ $(j = 1, 2, \cdots)$ with finite mean. Then, the total damage $Z_j \equiv \sum_{i=1}^{j} Y_i$ to the j-th damage shock where $Z_0 \equiv 0$ has a distribution

$$\begin{aligned} G^{(j)}(x) &\equiv \Pr\{Z_j \leq x\} \\ &= \begin{cases} 1 & (j = 0), \\ G_1 * G_2 * \cdots * G_j(x) & (j = 1, 2, \cdots), \end{cases} \end{aligned} \tag{2.2}$$

where the asterisk mark represents the Stieltjes convolution, *i.e.*, $a * b(t) \equiv \int_0^t b(t-u)da(u)$ for any functions $a(t)$ and $b(t)$.

Let $F(t) \equiv 1 - e^{-(1-p)\Gamma(t)}$, which is the distribution of failure time due to failure shock, and $\overline{F}(t) \equiv 1 - F(t)$. Then, the probability that a system is replaced before failure at

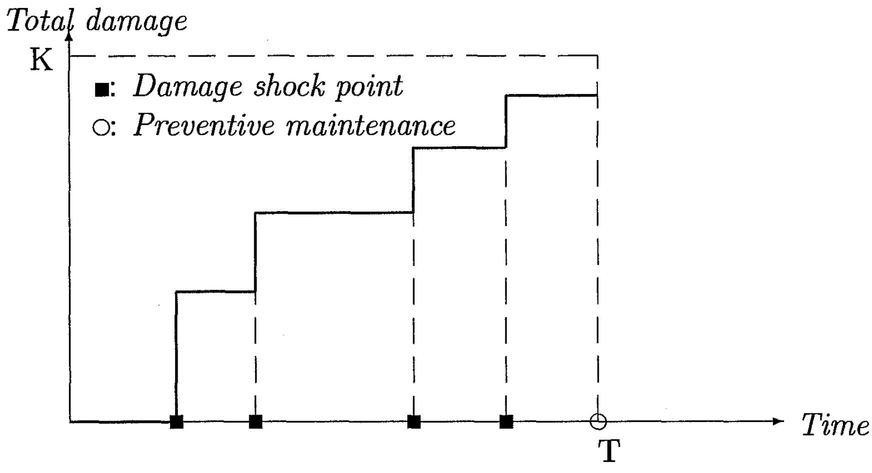

Figure 2: Preventive maintenance of cumulative damage model

scheduled time T is

$$\overline{F}(T)\sum_{j=0}^{\infty}H_j(T)G^{(j)}(K), \tag{2.3}$$

the probability that a system is replaced when the total damage has exceeded K is

$$\sum_{j=0}^{\infty}[G^{(j)}(K)-G^{(j+1)}(K)]\int_0^T \overline{F}(t)H_j(t)\lambda(t)dt, \tag{2.4}$$

and the probability that a system is replaced at failure shock is

$$\sum_{j=0}^{\infty}G^{(j)}(K)\int_0^T H_j(t)dF(t). \tag{2.5}$$

It is evident that (2.3)+(2.4)+(2.5)=1.

The mean time $E(T)$ to replacement is

$$\begin{aligned} E(T) &\equiv \sum_{j=0}^{\infty}[G^{(j)}(K)-G^{(j+1)}(K)]\int_0^T t\overline{F}(t)H_j(t)\lambda(t)dt \\ &\quad +\overline{F}(T)T\sum_{j=0}^{\infty}H_j(T)G^{(j)}(K)+\sum_{j=0}^{\infty}G^{(j)}(K)\int_0^T tH_j(t)dF(t) \\ &= \sum_{j=0}^{\infty}G^{(j)}(K)\int_0^T H_j(t)\overline{F}(t)dt. \end{aligned} \tag{2.6}$$

Further, the expected number of shocks until replacement is

$$\begin{aligned} &\sum_{j=0}^{\infty}(j+1)[G^{(j)}(K)-G^{(j+1)}(K)]\int_0^T \overline{F}(t)H_j(t)\lambda(t)dt \\ &+\overline{F}(T)\sum_{j=0}^{\infty}jH_j(T)G^{(j)}(K)+\sum_{j=0}^{\infty}jG^{(j)}(K)\int_0^T H_j(t)dF(t) \\ &= \sum_{j=0}^{\infty}G^{(j)}(K)\int_0^T H_j(t)\lambda(t)\overline{F}(t)dt. \end{aligned} \tag{2.7}$$

3. Backup Policy

In recent years, the database in computer systems has become very important in the highly information-oriented society. In particular, the reliable database is the most indispensable instrument in on-line transaction processing systems such as real-time systems used for account of bank. The data in a computer system are frequently updated by adding or deleting them, and are stored in floppy disks or other secondary media. However, data files in secondary media are sometimes broken by several errors due to noises, human errors and hardware faults. In this case, we have to reconstruct the same files from the beginning. The most simple and dependable method to ensure the safety of data would be always to make the backup copies of all files in other places, and to take out them if files in the original secondary media are broken. But, this method would take hours and costs, when files become large. To make the backup copies efficiently, we might dump only files which have changed since the last backup. This would reduce significantly both duration time and size of backup [16].

The recovery techniques for database failures [2,13], the backup policies for hard disks [14] and the checkpoint intervals [4, 7, 8, 20] were studied in many papers . In this paper, we apply the cumulative damage model to the backup of files for database media failures, by putting *damage shock* by *update*, *failure shock* by *database failure* and *damage* by *dumped files* : A database is updated at a nonhomogeneous Poisson process with an intensity function $\lambda(t) = p\gamma(t)$ and only files, which have changed or are new since the last backup, are dumped, which is called *incremental backup*. Further, the *full backup* is done at a specified day, and all files are dumped, *e.g.*, on the weekend, because it needs both enlarged time and size of backup. Suppose that a database system fails according to a distribution $F(t) = 1 - e^{-(1-p)\Gamma(t)}$.

To ensure the safety of data and to save hours, we make the following backup policy: If the total dumped files do not exceed a threshold level K, we perform the incremental backup where only new files since the previous full backup are dumped. Conversely, we perform the full backup at periodic time T, when the total files have exceeded K, or when the database fails, whichever occurs first. It is assumed that the database system returns to an initial state by the full backup.

Let introduce the following costs: A cost c_1 is suffered for the incremental backup, a cost $c_2 + c_0(x)$ is suffered for the full backup at time T when the total files are x $(0 \le x < K)$, a cost $c_3 + c_0(K)$ is suffered for the full backup when the total files have exceeded a threshold level K, and a cost $c_4 + c_0(x)$ is suffered for the recovery when the database fails, where $c_1 \le c_2 < c_3 \le c_4$ and $c_0(0) \equiv 0$. Then, from (2.3), (2.4), (2.5) and (2.7), the expected cost to full backup is

$$\begin{aligned} & c_1 \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt + \overline{F}(T) \sum_{j=0}^{\infty} H_j(T) \int_0^K [c_2 + c_0(x)]dG^{(j)}(x) \\ + \; & [c_3 + c_0(K)] \sum_{j=0}^{\infty} [G^{(j)}(K) - G^{(j+1)}(K)] \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ + \; & \sum_{j=0}^{\infty} \int_0^T H_j(t)dF(t) \int_0^K [c_4 + c_0(x)]dG^{(j)}(x). \end{aligned} \tag{3.1}$$

Therefore, from (2.6) and (3.1), the expected cost per unit of time is

$$C(T) = A(T)/E(T), \tag{3.2}$$

where $E(T)$ is given in (2.6), and

$$\begin{aligned} A(T) \equiv\ & c_1 \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & + c_2 \overline{F}(T) \sum_{j=0}^{\infty} H_j(T) G^{(j)}(K) \\ & + c_3 \sum_{j=0}^{\infty} [G^{(j)}(K) - G^{(j+1)}(K)] \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & + c_4 \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)dF(t) \\ & + \sum_{j=0}^{\infty} \int_0^K [1 - G^{(j)}(x)]dc_0(x) \int_0^T \overline{F}(t)dH_j(t). \end{aligned} \tag{3.3}$$

4. Optimal Policy

Suppose that $c_0(x) = c_0 x$, *i.e.*, the proportional cost of full backup is a linear function of dumped files, because the full backup cost would increase in proportion to quantities of floppy disks or other storage files, and usage times spent for dumped files. In this case, the numerator $A(T)$ in (3.3) is rewritten as

$$\begin{aligned} A(T) =\ & [c_1 + (c_4 - c_2)(1-p)/p] \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & + c_2 + (c_3 - c_2) \sum_{j=0}^{\infty} [G^{(j)}(K) - G^{(j+1)}(K)] \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & + c_0 \sum_{j=0}^{\infty} \int_0^K [G^{(j)}(x) - G^{(j+1)}(x)]dx \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt. \end{aligned} \tag{4.1}$$

Thus, if $M(K) \equiv \sum_{j=1}^{\infty} G^{(j)}(K) < \infty$, then, $C(0) \equiv \lim_{T\to 0} C(T) = \infty$, and hence, there exists a positive T^* $(0 < T^* \leq \infty)$ which minimizes $C(T)$.

A necessary condition that a finite T^* minimizes $C(T)$ is given by differentiating $C(T)$ with respect to T and setting it equal to zero. Hence, from (3.2) and (4.1), we have

$$\begin{aligned} & U(T) \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\overline{F}(t)dt \\ & -[c_1 + (c_4 - c_2)(1-p)/p] \sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & -(c_3 - c_2) \sum_{j=0}^{\infty} [G^{(j)}(K) - G^{(j+1)}(K)] \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt \\ & -c_0 \sum_{j=0}^{\infty} \int_0^K [G^{(j)}(x) - G^{(j+1)}(x)]dx \int_0^T H_j(t)\lambda(t)\overline{F}(t)dt = c_2, \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} U(T) \equiv\ & [c_1 + (c_4 - c_2)(1-p)/p]\lambda(T) + \{(c_3 - c_2) \sum_{j=0}^{\infty} [G^{(j)}(K) - G^{(j+1)}(K)] H_j(T) \\ & + c_0 \sum_{j=0}^{\infty} \int_0^K [G^{(j)}(x) - G^{(j+1)}(x)]dx H_j(T)\}\lambda(T) / \sum_{j=0}^{\infty} G^{(j)}(K) H_j(T). \end{aligned} \tag{4.3}$$

Letting $L(T)$ be the left-hand side of (4.2),

$$\begin{aligned} L(0) &\equiv \lim_{T\to 0} L(T) = 0, \\ L(\infty) &\equiv \lim_{T\to\infty} L(T) \\ &= U(\infty)E(\infty) - [c_1 p + (c_4 - c_2)(1-p)]\sum_{j=0}^{\infty} p^j G^{(j)}(K) \\ &\quad -(c_3 - c_2)\sum_{j=0}^{\infty} p^{j+1}[G^{(j)}(K) - G^{(j+1)}(K)] \\ &\quad -c_0 \sum_{j=0}^{\infty} p^{j+1} \int_0^K [G^{(j)}(x) - G^{(j+1)}(x)]dx, \qquad (4.4) \\ L'(T) &= U'(T)\sum_{j=0}^{\infty} G^{(j)}(K) \int_0^T H_j(t)\overline{F}(t)dt, \end{aligned}$$

where $E(\infty) \equiv \lim_{T\to\infty} E(T)$ and $U(\infty) \equiv \lim_{T\to\infty} U(T)$.

If $U'(T) > 0$ and $L(\infty) > c_2$, then $L(T)$ is a strictly increasing function from 0 to $L(\infty)$, and hence, there exists a finite and unique T^* $(0 < T^* < \infty)$ which satisfies (4.2), and the resulting cost is

$$C(T^*) = U(T^*). \qquad (4.5)$$

Conversely, if $U'(T) > 0$ and $L(\infty) \leq c_2$ or $U'(T) \leq 0$ then $T^* = \infty$.

In particular, note that $L(\infty) = \infty$ when the database is updated at a nonhomogeneous Poisson process with a mean-value function $R(t) \equiv p\lambda t^m$ $(m > 1)$ [18].

5. Numerical Examples

Suppose that the database is updated at a Poisson process with rate $p\lambda$, *i.e.*, $\lambda(t) = p\lambda$, $R(t) = p\lambda t$, $H_j(t) = [(p\lambda t)^j/j!]e^{-p\lambda t}$ $(j = 0, 1, 2, \cdots)$ and $F(t) = 1 - e^{-(1-p)\lambda t}$.

5.1. Exponential case

We compute the optimal policy numerically when $G_j(x) = 1 - e^{-\mu x}$, *i.e.*, $G^{(j)}(x) = 1 - \sum_{i=0}^{j-1}[(\mu x)^i/i!]e^{-\mu x}$ $(j = 1, 2, \cdots)$ and $M(K) = \mu K$. In this case, when $c_3 - c_2 - c_0/\mu \neq 0$, equation (4.2) is

$$\sum_{j=0}^{\infty}[V(T)G^{(j)}(K) - \frac{(\mu K)^j}{j!}e^{-\mu K}] \int_0^T H_j(t) p\lambda e^{-(1-p)\lambda t} dt = \frac{c_2}{c_3 - c_2 - c_0/\mu}, \qquad (5.1)$$

where

$$V(T) \equiv \frac{\sum_{j=0}^{\infty} H_j(T)[G^{(j)}(K) - G^{(j+1)}(K)]}{\sum_{j=0}^{\infty} H_j(T) G^{(j)}(K)}. \qquad (5.2)$$

Note that $G^{(j+1)}(K)/G^{(j)}(K)$ is strictly decreasing in j when $G_j(x) = 1 - e^{-\mu x}$. Thus, $V(T)$ is strictly increasing in T [15], and $V(\infty) \equiv \lim_{T\to\infty} V(T) = 1$. Letting $Q(T)$ be the left-hand side of (5.1), it is evident that

$$\begin{aligned} Q(0) &\equiv \lim_{T\to 0} Q(T) = 0, \\ Q(\infty) &\equiv \lim_{T\to\infty} Q(T) = \sum_{j=1}^{\infty} p^j G^{(j)}(K) \end{aligned}$$

$$= \begin{cases} \frac{p}{1-p}[1-e^{-(1-p)\mu K}] & (p<1), \\ \mu K & (p=1), \end{cases} \tag{5.3}$$

$$Q'(T) = V'(T)\sum_{j=0}^{\infty} G^{(j)}(K)\int_0^T H_j(t)p\lambda e^{-(1-p)\lambda t}dt > 0.$$

Thus, if $p<1$, then we have the following optimal policy:

(i) If $c_3-c_2-c_0/\mu>0$ and $\frac{p}{1-p}(1-e^{-(1-p)\mu K}) > \frac{c_2}{c_3-c_2-c_0/\mu}$ then there exists a finite and unique T^* $(0<T^*<\infty)$ which satisfies (5.1), and the resulting cost is

$$\frac{C(T^*)}{\lambda} = pc_1 + (1-p)(c_4-c_2) + \frac{pc_0}{\mu} + p(c_3-c_2-\frac{c_0}{\mu})V(T^*). \tag{5.4}$$

(ii) If$c_3-c_2-c_0/\mu \le 0$ or $\frac{p}{1-p}(1-e^{-(1-p)\mu K}) \le \frac{c_2}{c_3-c_2-c_0/\mu}$ then $T^*=\infty$, and the resulting cost is

$$\frac{C(\infty)}{\lambda} = pc_1 + (1-p)c_4 + \frac{pc_0}{\mu} + \frac{p(1-p)}{e^{(1-p)\mu K}-p}(c_3-\frac{c_0}{\mu}). \tag{5.5}$$

It is easily seen that $Q(T)$ is strictly increasing in p since $V(T)$ is also strictly increasing in p. Hence, optimal T^* in case(i) is a decreasing function of p.

In particular, when $p=1$, the optimal policy is rewritten as:

(iii) If $c_3-c_2-c_0/\mu>0$ and $\mu K > \frac{c_2}{c_3-c_2-c_0/\mu}$ then there exists a finite and unique T^* which satisfies (5.1), and the resulting cost is

$$\frac{C(T^*)}{\lambda} = c_1 + \frac{c_0}{\mu} + (c_3-c_2-\frac{c_0}{\mu})V(T^*). \tag{5.6}$$

(iv) If $c_3-c_2-c_0/\mu \le 0$ or $\mu K \le \frac{c_2}{c_3-c_2-c_0/\mu}$ then $T^*=\infty$, and the resulting cost is

$$\frac{C(\infty)}{\lambda} = c_1 + \frac{c_3+c_0K}{1+\mu K}. \tag{5.7}$$

Table 1 Optimal full backup time λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ when $c_1/c_2=0.5$, $c_3/c_2=4$, $c_4/c_2=25$ and $c_0/(c_2\mu)=0.1$

μK	p					
	1.00		0.98		0.96	
	λT^*	$C(T^*)/(\lambda c_2)$	λT^*	$C(T^*)/(\lambda c_2)$	λT^*	$C(T^*)/(\lambda c_2)$
8	5.365	0.908	5.566	1.377	5.784	1.447
10	6.459	0.835	6.712	1.307	6.986	1.378
12	7.627	0.788	7.937	1.260	8.275	1.332
14	8.848	0.755	9.222	1.228	9.629	1.302
16	10.111	0.731	10.554	1.205	11.037	1.280
18	11.410	0.713	11.926	1.188	12.490	1.262

Table 1 gives the optimal full backup times λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ for $p=1.00, 0.98, 0.96$ and $\mu K = 8, 10, 12, 14, 16, 18$ when $c_1/c_2=0.5$, $c_3/c_2=4$, $c_4/c_2=25$ and $c_0/(c_2\mu)=0.1$. Note that all costs are relative to cost c_2 and all times are relative to $1/\lambda$. Similarly, Table 2 gives the optimal full backup times λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ for $c_3/c_2=3, 6, 12, 24$ and $c_0/(c_2\mu)=0.01, 0.10, 1.00$ when $c_1/c_2=0.5$, $c_4/c_2=$

Table 2 Optimal full backup time λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ when $c_1/c_2 = 0.5$, $c_4/c_2 = 25$, $\mu K = 12$ and $p = 0.98$

c_3/c_2	$c_0/(c_2\mu)$					
	0.01		0.10		1.00	
	λT^*	$C(T^*)/(\lambda c_2)$	λT^*	$C(T^*)/(\lambda c_2)$	λT^*	$C(T^*)/(\lambda c_2)$
3	9.086	1.154	9.225	1.239	12.015	2.093
6	6.656	1.202	6.694	1.289	7.138	2.159
12	5.264	1.253	5.277	1.340	5.409	2.216
24	4.285	1.310	4.290	1.399	4.337	2.277

25, $p = 0.98$ and $\mu K = 12$. In this case, from optimal policy (i), finite λT^* exist uniquely if $c_3/c_2 - c_0/(c_2\mu) > 1 + (1 - 0.98)/[0.98(1 - e^{(1-0.98)\times 12})] \approx 1.096$. These show that the optimal λT^* are increasing with both $c_0/(c_2\mu)$ and μK, and conversely, are decreasing when c_3/c_2 is increasing, and the costs $C(T^*)/(\lambda c_2)$ are increasing with both $c_0/(c_2\mu)$ and c_3/c_2, and conversely, are decreasing when μK is increasing. Further, both optimal times λT^* and resulting costs $C(T^*)/(\lambda c_2)$ become small as p becomes large, because the mean time $1/(p\lambda)$ of update becomes small.

For example, when $c_3/c_2 = 4$, $c_0/(c_2\mu) = 0.1$, $\mu K = 14$ and $p = 0.98$, the optimal full backup time λT^* is about 9.2. That is, when the mean time of update is $1/(p\lambda) = 1$ day, the optimal full backup time T^* is about $9.2p \approx 9$ days. Taking another point of view, we can intuitively see that a preventive full backup should be approximately made when the total dumped files have exceeded $(0.98 \times 9.2)/14 \approx 65\%$ of a threshold level K, since $\mu K = 14$ represents the expected number of updates until the total dumped files exceed K. Moreover, the expected number μK becomes large as $1/\mu$ is small, and in this case, the optimal times T^* also become large, and conversely, the resulting costs $C(T^*)/(\lambda c_2)$ become small.

5.2. Different exponential case

Next, suppose that the amount Y_j of newly dumped files at the j-th update has different exponential distributions, *i.e.*, $G_j(x) = 1 - e^{-\mu_j x}$ $(j = 1, 2, \cdots)$. We show that an amount Y_j of files which is dumped at the j-th update increases in a geometric ratio. Suppose that an amount of files at some update is Y, the total volume of files is M and the total files which have been already dumped is A $(0 \le A \le M)$. Then, we assume that an amount of newly dumped files is proportional to the vacant space, *i.e.*, $Y \times (M - A)/M$. From the above assumption, we have,

$$Y_{j+1} = \begin{cases} Y & (j = 0), \\ Y \times (M - \sum_{i=1}^{j} Y_i)/M & (j = 1, 2, \cdots). \end{cases} \tag{5.8}$$

Solving this equation,

$$Y_j = Y \times (1 - Y/M)^{j-1} \quad (j = 1, 2, \cdots). \tag{5.9}$$

We define that $Y/M \equiv 1 - \alpha$ which is an amount ratio of dumped files at the first update. Then, $Y_j/M = \alpha^{j-1}(1 - \alpha)$ $(j = 1, 2, \ldots)$ which is a geometric distribution with mean $1/(1-\alpha)$. This shows that an amount of newly dumped files forms a geometric process with Y_1/a^{j-1} $(j = 1, 2, \cdots)$ where $1/a = \alpha$ in [19].

In particular, when Y_j increases in a geometric ratio, *i.e.*, $Y_j = \alpha^{j-1}Y$ and $1/\mu_j \equiv \alpha^{j-1}/\mu$ $(0 < \alpha < 1)$. Then, the distribution of total files until the j-th update is easily given by

$$G^{(j)}(x) = \begin{cases} 1 - e^{-\mu x} & (j = 1), \\ 1 - \sum_{l=1}^{j} \left[\prod_{i=1, i\neq l}^{j} \frac{1}{1-\alpha^{i-l}}\right] e^{-\frac{\mu x}{\alpha^{l-1}}} & (j = 2, 3, \cdots). \end{cases} \tag{5.10}$$

Then, using the relation from Appendix A,

$$\int_0^K [G^{(j-1)}(x) - G^{(j)}(x)]dx = \frac{1}{\mu_j} G^{(j)}(K), \tag{5.11}$$

equation (4.2) is

$$\sum_{j=0}^{\infty} [(c_3 - c_2 - c_0\alpha^j/\mu) G^{(j+1)}(K) - W(T) G^{(j)}(K)] \int_0^T H_j(t) p\lambda e^{-(1-p)\lambda t} dt = c_2, \tag{5.12}$$

and equation (4.5) is

$$C(T^*)/\lambda = p(c_1 + c_3) + (1-p)c_4 - c_2 - pW(T^*), \tag{5.13}$$

where

$$W(T) \equiv \frac{\sum_{j=0}^{\infty} (c_3 - c_2 - c_0\alpha^j/\mu) H_j(T) G^{(j+1)}(K)}{\sum_{j=0}^{\infty} H_j(T) G^{(j)}(K)}. \tag{5.14}$$

Table 3 gives the optimal full backup times λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ for $c_3/c_2 = 5, 10, 20$ and $\alpha = 1.00, 0.95, 0.90, 0.85, 0.80, 0.75$ when $p = 0.98$, $c_1/c_2 = 0.5$, $c_4/c_2 = 25$, $c_0/(c_2\mu) = 0.1$ and $\mu K = 12$. This shows that the optimal times λT^* are increasing when c_3/c_2 and α are decreasing, and conversely, the costs $C(T^*)/(\lambda c_2)$ are decreasing with both c_3/c_2 and α. However, they are almost unchanged for α.

Table 3 Optimal full backup time λT^* and the resulting costs $C(T^*)/(\lambda c_2)$ when $\mu K = 12$, $p = 0.98$, $c_1/c_2 = 0.5$, $c_4/c_2 = 25$, and $c_0/(c_2\mu) = 0.1$

α	c_3/c_2					
	5		10		20	
	λT^*	$C(T^*)/(\lambda c_2)$	λT^*	$C(T^*)/\lambda c_2)$	λT^*	$C(T^*)/(\lambda c_2)$
1.00	7.179	1.276	5.594	1.326	4.522	1.382
0.95	7.186	1.268	5.609	1.318	4.541	1.374
0.90	7.244	1.259	5.650	1.309	4.569	1.366
0.85	7.314	1.249	5.691	1.302	4.594	1.359
0.80	7.352	1.242	5.717	1.294	4.613	1.353
0.75	7.363	1.235	5.732	1.288	4.625	1.347

6. Conclusions

We have proposed the extended cumulative damage model with two kinds of shocks where a system fails or suffers only damage, and is replaced at scheduled time T. Using the theory of cumulative processes, we derive the expected cost and discuss the optimal replacement policy which minimizes it.

Further, we have shown that this would be applied to the backup of secondary storage files in the database system. Thus, by estimating the costs of backups and the amount of dumped files from actual data and by modifying some suppositions, we could practically determine a scheduled time of full backup. These formulations and results would be applied to other management policies for computer systems [15].

Acknowledgment This form a part of research results by the Hori Information Science Promotion Foundation.

Appendix A
Derivation of equation (5.11)

Let $g_j(s)$ denotes the Laplace-Stiltjes (LS) transform of Cdf $G_j(x)$, *i.e.*,

$$g_j(s) \equiv \int_0^\infty e^{-sx} dG_j(x), \tag{A.1}$$

for $s > 0$, and $g^{(j)}(s)$ denotes the Laplace-Stiltjes (LS) transform of Cdf $G^{(j)}(x)$. Then, we easily have

$$g^{(j)}(s) = g^{(j-1)}(s) g_j(s), \tag{A.2}$$

and

$$g_j(s) = \frac{\mu_j}{s + \mu_j}, \tag{A.3}$$

when $G_j(x) = 1 - e^{-\mu_j}$ $(j = 0, 1, 2, \cdots)$. Thus,

$$\frac{1}{s}[g^{(j-1)}(s) - g^{(j)}(s)] = \frac{1}{\mu_j} g^{(j)}(s). \tag{A.4}$$

Inverting the LS transforms of (A.4), we have

$$\int_0^K [G^{(j-1)}(x) - G^{(j)}(x)] dx = \frac{1}{\mu_j} G^{(j)}(K). \tag{A.5}$$

References

[1] R.E. Barlow and F. Proschan: *Mathematical Theory of Reliability* (John Wiley & Sons, New York, 1965).

[2] K. M. Chandy, J. C. Browne, C. W. Dissly and W. R. Uhrig: Analytic models for rollback and recovery strategies in data base systems. *IEEE Trans. Software Engineering,* **SE-1**(1975)100-110.

[3] D.R. Cox: *Renewal Theory* (Methuen, London, 1962).

[4] T. Dohi, T. Aoki, N. Kaio and S. Osaki: Computational aspects of optimal checkpoint strategy in fault-tolerant database management. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences,* **E80-A**(1997)2006-2015.

[5] J.D. Esary, A.W. Marshall and F. Proschan: Shock models and wear processes. *Annals of Probability,* **1**(1973) 627-649.

[6] R.M. Feldman: Optimal replacement with semi-Markov shock models. *Journal of Applied Probability,* **13**(1976) 108-117.

[7] S. Fukumoto, N. Kaio and S. Osaki: A study of checkpoint generations for a database recovery mechanism. *Computers & Mathematics with Applications,* **1**(1992)63-68.

[8] E. Gelenbe: On the optimum checkpoint interval.*Journal of ACM,* **26**(1979)259-270.

[9] T. Nakagawa: On a replacement problem of a cumulative damage model. *Operational Research Quarterly,* **27**(1976) 895-900.

[10] T. Nakagawa: A summary of discrete replacement policies. *European Journal of Operational Research,* **17**(1984) 382-392.

[11] T. Nakagawa and M. Kijima: Replacement policies for a cumulative damage model with minimal repair at failure. *IEEE Trans. Reliability,* **13**(1989) 581-584.

[12] S. Osaki: *Applied Stochastic System Modeling* (Springer Verlag, Berlin, 1992).

[13] A. Reuter: Performance analysis of recovery techniques. *ACM Trans. Database Systems,* **4**(1984)526-559.

[14] H. Sandoh, N. Kaio and H. Kawai: On backup policies for computer disks. *Reliability Engineering & System Safety,* **37**(1992)29-32.

[15] T. Satow, K. Yasui and T. Nakagawa: Optimal garbage collection policies for a database in a computer system. *RAIRO-Operations Research,* **30** (1996)359-372.

[16] K. Suzuki and K. Nakajima: Storage management software. *Fujitsu,* **48**(1993) 389-397.

[17] H.M. Taylor: Optimal replacement under additive damage and other failure models. *Naval Research Logistics Quarterly,* **22**(1975) 1-18.

[18] S. Yamada: *Software Reliability Models-Foundation and Application* (JUSE Press Ltd, Tokyo, 1994)82-84.

[19] Lam Yeh: Optimal geometric process eplacement model.*Acta Mathematicae Applicatae Sinica,* **8**(1992)73-81.

[20] J. W. Young: A first order approximation to the optimum checkpoint interval. *Communications of ACM,* **17**(1974)530-531.

Toshio Nakagawa
Department of Industrial Engineering
Aichi Institute of Technology
1247 Yachigusa, Yagusa-cho
Toyota 470-0392 JAPAN
E-mail: nakagawa@ie.aitech.ac.jp