Journal of the Operations Research
Society of Japan
Vol. 35, No. 4, December 1992

AN OPTIMAL 1/N BACKUP POLICY FOR DATA FLOPPY DISKS UNDER EFFICIENCY BASIS

Hiroaki Sandoh
University of Marketing & Distribution Sciences

Hajime Kawai
Tottori University

(Received September 30, 1991)

Abstract A word processor has become one of essential devices used in offices. Document files created by using a word processor are generally preserved on a floppy disk or floppy disks for references. Nevertheless, the files stored on a floppy disk are occasionally lost due to human errors, the life of the floppy disk and a failure of hardware devices, which comprise the word processor. This is called *a floppy disk failure*. One of the simplest methods for protecting us from such serious losses is to backup files on another floppy disk.

Frequent backup operations would spend much time in themselves although they could reduce the loss at a floppy disk failure. On the contrary, occasional backup operations would make the loss at a floppy disk failure very large although they could save time in backup operations. These observations reveal the significance of determining an adequate backup timing of files.

The present study proposes a $1/N$ backup policy for data floppy disks, which suggests to backup files when $1/N$ of the total memory of a floppy disk is consumed. The availability of the proposed backup policy is formulated as an objective function under efficiency basis. The existence of an optimal integer N^* that maximizes the availability is then examined to clarify the conditions under which such an optimal integer exists. Numerical examples are also presented.

1 Introduction

Word processors are recently used in offices very widely. Document files created and/or updated by using a word processor are generally preserved on a floppy disk or floppy disks. However, the files on a floppy disk are occasionally lost because of human errors, the life of the floppy disk or a failure of hardware devices that comprise a word processor. This is called *a floppy disk failure*. Since the value of information has recently increased, a floppy disk failure is a serious loss in many cases. It is important to backup files on another floppy disk periodically to protect their users from such a loss. In the case of a data floppy disk failure, its backup disk can partially recover the original data floppy disk. The recovery may be partial since the backup disk only stores the files at the last backup operation.

It would spend much time in backup operations to backup files frequently although it could reduce the loss at a floppy disk failure. On the contrary, rare backup operations would make the loss at a floppy disk failure very large, while the time spent in backup operations could sharply be reduced. These observations indicates the significance of determining the adequate timing of backup operations.

Similar problems can be observed in determining the backup times for files stored in a hard computer disk that is used for a personal computer or an engineering work station. For such problems, Sandoh and Kawai [1] have proposed an N-job backup policy, which suggests to backup files when N jobs updating and/or creating files are finished. The availability was adopted as an objective function to be maximized. Sandoh, Kawai and Ibaraki [2] have also discussed another backup policy, where files are backed up at their age T. The availability was also formulated as an objective function to be maximized.

This study proposes a backup policy peculiar to data floppy disks, where files are backed up when $1/N$ of the total memory of a floppy disk is consumed. This is called a $1/N$ *backup policy.* The availability of this policy is formulated as an objective function under efficiency basis. The design variable of this policy is an integer N, and an integer $N = N^*$ is optimal if N^* maximizes the availability. The existence of such an optimal integer is then analyzed to clarify the conditions under which N^* exists. Numerical examples are also presented to illustrate the theoretical underpinnings of the proposed backup policy formulation.

2 Availability

Consider a backup policy which suggests to backup files on another floppy disk when $1/N$ of the memory of a data floppy disk is consumed. We also consider that at each backup time, only the files updated and/or created since the previous backup operation are backed up.

Let us assume that the floppy disk failure time, X, follows an exponential distribution with a failure rate, λ, considerng that they seldom occur. It is also assumed that a floppy disk failure can instantly be detected, and that any floppy disk failure does not occur during a recovery operation although one may occur during a backup operation.

At each backup time, the backing up time consists of a setup time, $\tau(\geq 0)$ and the time, a/N, which is proportional to the memory newly consumed,where $a(> 0)$ denotes the proportional constant. In the following, we regard the time required for consuming all the memory of a data floppy disk as a unit time.

Let us here define the availability as

$$W(N) = \lim_{t\to+\infty} \frac{E[\text{the total effective time on}(0,t)]}{t}, \tag{1}$$

where the effective time signifies the time during which creation and/or updating of files were followed by a successful backup operation.

Under the above assumptions, a renewal process [3] is generated, where one of the following two events occurs corresponds to a renewal point:

(a) A backup operation has successfully been completed.

(b) Following a floppy disk failure, a recovery operation by using backup disks has just been ended. It should be noticed that at this time, the contents of the data floppy disk agree with those at the completion of the last backup operation.

By the renewal theory [3], the availability defined in Eq. (1) can be rewritten as

$$W(N) = \frac{E[\text{the effective time of 1 cycle}]}{E[\text{the length of 1 cycle}]}, \tag{2}$$

where 1 cycle refers to the interval between two successive renewal points.

When the event (a) occurs, the effective time of 1 cycle is $1/N$, while the length of 1 cycle is $(a+1)/N+\tau$. In the case of the event (b), there does exist no effective time since all the information updated after the last backup operation are lost, and the length of 1 cycle is $X+Y$. Figure 1 shows the effective time, the ineffective time and the length of 1 cycle in each case of the events (a) and (b).

Let $A(N)$ and $B(N)$ respectively denote the denominator and numerator of the right-hand-side of Eq. (2), i. e., let

$$W(N) = \frac{B(N)}{A(N)}, \tag{3}$$

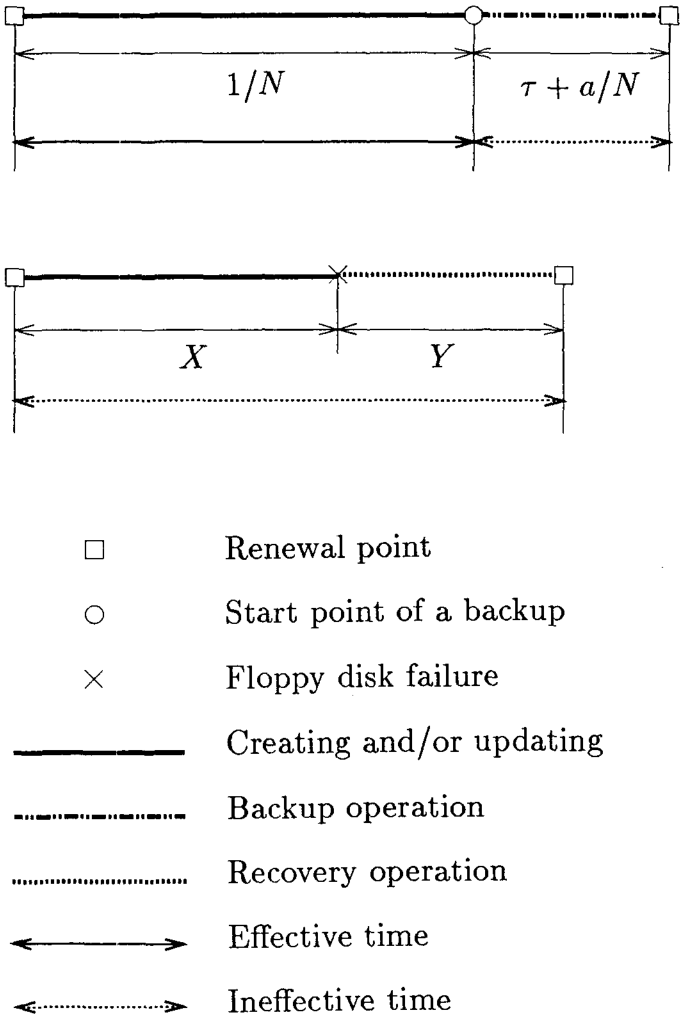

Figure 1: 1 cycle of the process.

then we have

$$A(N) = (\frac{a+1}{N} + \tau)\overline{F}(\frac{a+1}{N} + \tau) + \int_0^\infty [\int_0^{\frac{a+1}{N}+r} (x+y)dF(x)]dG(y), \tag{4}$$

$$B(N) = \frac{1}{N}\overline{F}(\frac{a+1}{N} + r), \tag{5}$$

where $F(x)$ denotes the cumulative distribution function (cdf) of X and

$$\overline{F}(x) = 1 - F(x), \tag{6}$$

and Y is a random variable that denotes the recovery time with cdf $G(y)$. The first and the second terms of the right-hand-side of Eq. (4) correspond to the cases of the events (a) and (b), respectively.

Since we assume that

$$F(x) = 1 - e^{-\lambda}, \tag{7}$$

$A(N)$ and $B(N)$ in Eqs. (4) and (5) become

$$A(N) = (\frac{1}{\lambda} + \mu)[1 - e^{-\lambda(\frac{a+1}{N}+r)}], \tag{8}$$

$$B(N) = \frac{1}{N}e^{-\lambda(\frac{a+1}{N}+r)}, \tag{9}$$

where

$$\mu = E[Y] = \int_0^\infty y dG(y), \tag{10}$$

which means the mean recovery time from a floppy disk failure. From Eqs. (8) and (9), the availability $W(N)$ in Eq. (3) is given by

$$W(N) = \frac{1}{(\frac{1}{\lambda} + \mu)N[e^{\lambda(\frac{a+1}{N}+r)} - 1]}. \tag{11}$$

We have formulated the availability $W(N)$. If an integer $N = N^*$ maximizes $W(N)$, then it is optimum. The next section examines the existence of such an optimal integer N^*.

3 Optimal Backup Policy

It is obvious that maximization of $W(N)$ with respect to N agrees with minimization of $D(N)$, where $D(N)$ is defined as

$$D(N) = N[e^{\lambda(\frac{a+1}{N}+r)} - 1]. \tag{12}$$

This implies that the optimal integer N^* is not affected by mean recovery time, μ, since $D(N)$ does not include μ.

Let us define u as

$$u = 1/N, \tag{13}$$

then u satisfies

$$u \in (0, 1], \tag{14}$$

and D(N) becomes

$$D(u) = \frac{e^{\lambda[(a+1)u+r]} - 1}{u}. \tag{15}$$

In the following, we consider to minimize $D(u)$ in relation to u.

By differntiating $D(u)$ with respect to u, we obtain

$$D'(u) = \frac{1}{u^2}\{e^{\lambda[(a+1)u+\tau]}[\lambda(a+1)u-1]+1\}. \tag{16}$$

The sign of $D'(u)$ agress with that of $L(u)$ defined as

$$L(u) = e^{\lambda[(a+1)u+\tau]}[\lambda(a+1)u-1]+1. \tag{17}$$

Then we have

$$\lim_{u\to+0} L(u) = -e^{\lambda\tau}+1 \begin{cases} =0, \tau=0, \\ <0, \tau>0, \end{cases} \tag{18}$$

$$L(1) = e^{\lambda(a+1+\tau)}[\lambda(a+1)-1]+1. \tag{19}$$

We further have

$$L'(u) = [\lambda(a+1)]^2 u e^{\lambda[(a+1)u+\tau]} > 0, \tag{20}$$

which reveals that $L(u)$ is increasing in $u \in (0,1]$.

Based on the above results, the existence of an optimal integer can be discussed for the fllowing two cases:

(1) $\tau = 0$ (i. e., the setup times is negligibly small):

In this case, Eqs. (18) and (20) indicate that $L(u) > 0$, i. e., $D'(u) > 0$ for $u > 0$, and thus $D(u)$ is increasing in u. This result recommends us to backup files as frequently as possible although there does not explicitly exist an optimal solution.

(2) $\tau > 0$ (i. e., the setup time cannot be neglected):

In this case, the existence of an optimal solution can furthermore be discussed for the following two subcases;

(a) $e^{\lambda(a+\tau+1)}[\lambda(a+1)-1]+1 > 0$.

From Eqs. (18), (19) and (20), the sign of $L(u)$ changes from negative to positive. It follows that $D(u)$ first decreases and then increases on (0,1] and that there exists a unique optimal solution u^*, which minimizes $D(u)$. This result implies that the optimal integer is $N^* = [1/u^*]$ if $W([1/u^*]) \geq W([1/u^*]+1)$, otherwise $N^* = [1/u^*]+1$, where $[\cdot]$ signifies the largest integer which does not exceed $\cdot$.

(b) $e^{\lambda(a+\tau+1)}[\lambda(a+1)-1]+1 \leq 0$.

Equations (18), (19) and (20) reveal that $L(u) \leq 0$ for $u > 0$ and that $D(u)$ decreases with u. The optimal solution is, therefore, $u^* = 1$, i. e., $N^* = 1$, which suggests to backup files when all the memory of a data floppy disk is just consumed.

4 Numerical Examples

This section presents numerical examples to illustrate the theoretical underpinnings of the proposed backup policy formulation.

Figure 2 reveals the availability functions for $\tau = 0.001, 0.005, 0.01$ in the case of (a, μ, λ)

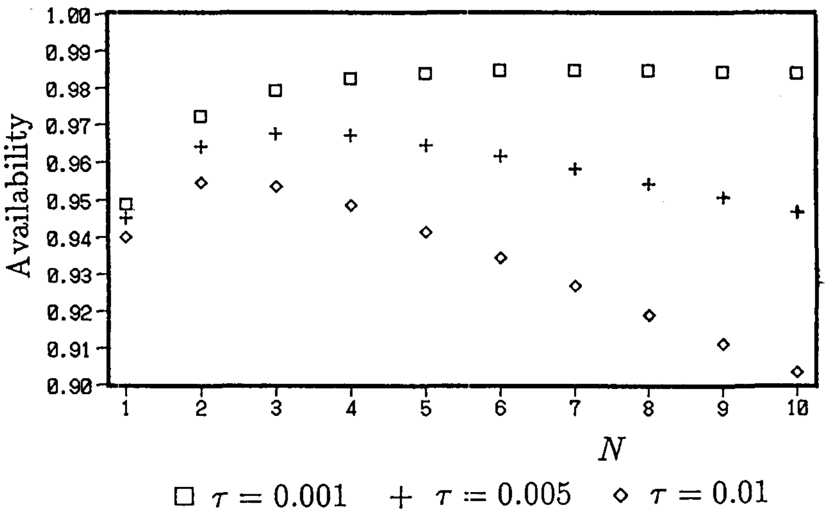

Figure 2: Availability.

Table 1: Optimal backup times.

$(a, \mu) = (0.001, 0.01)$

λ	$\tau = 0.001$		$\tau = 0.005$		$\tau = 0.01$	
	N^*	$W(N^*)$	N^*	$W(N^*)$	N^*	$W(N^*)$
0.01	2	0.9945	1	0.9890	1	0.9840
0.02	3	0.9929	2	0.9841	1	0.9792
0.03	4	0.9913	2	0.9816	1	0.9742
0.04	5	0.9900	2	0.9792	2	0.9695
0.05	5	0.9890	2	0.9767	2	0.9670
0.06	6	0.9881	3	0.9743	2	0.9645
0.07	6	0.9872	3	0.9726	2	0.9620
0.08	6	0.9864	3	0.9710	2	0.9596
0.09	7	0.9856	3	0.9693	2	0.9571
0.1	7	0.9849	3	0.9677	2	0.9546
0.2	10	0.9792	5	0.9548	3	0.9370
0.3	12	0.9747	6	0.9451	4	0.9236
0.4	14	0.9710	6	0.9370	5	0.9120
0.5	16	0.9677	7	0.9300	5	0.9024
0.6	18	0.9648	8	0.9236	6	0.8934
0.7	19	0.9621	9	0.9177	6	0.8854
0.8	20	0.9596	9	0.9123	7	0.8777
0.9	22	0.9572	10	0.9072	7	0.8709
1.0	23	0.9550	10	0.9024	7	0.8641

= (0.001, 0.01, 0.1). It is observed in Fig. 2 that the optimal backup time becomes important as the setup time increases. When a is set to 0.0025 and 0.005, quite similar results to those in Fig. 2 were obtained except the fact that the availability slightly decreases with increasing a.

Table 1 indicates the optimal integer N^* and its corresponding availability for $(a, \mu) = (0.001, 0.01)$ when τ and λ changes. It is seen in Table 1 that the optimal integer increases with λ, which suggests to backup files more frequently as λ becomes large.

5 Conclusions

This study proposed an optimal $1/N$ backup policy peculiar to data floppy disks, which suggests to backup files on a floppy disk when $1/N$ of the total memory of a floppy disk is consumed. The availability of the proposed policy is formulated with a view to determining the optimal integer, N^*, under efficiency basis. The existence of such an optimal integer is also analyzed to clarify the conditions under which N^* exists. Numerical examples are presented to illustrate the proposed backup policy.

Acknowledgement

This research was supported in part by Grant-in-Aids (03780040) for Scientific Research of the Ministry of Education, Science and Culture.

References

[1] Sandoh, H. and Kawai, H.: An Optimal N-Job Backup Policy Maximizing Availability for a Hard Computer Disk. *Journal of the Operations Research Society of Japan*, Vol. 33, No. 4, pp. 383–390 1991.

[2] Sandoh, H., Kawai, H. and Ibaraki, T.: An Optimal Backup Policy for a Hard Computer Disk Depending on Age under Availability Criterion. *Computers & Mathematics with Applications*, Vol. 24, No. 1/2, pp. 57–62 1992.

[3] Ross, S. M.: *Applied Probability Models with Optimization Applications*. Holden-Day, San Francisco, 1970.

Hiroaki SANDOH:
Department of Information &
Management Sciences
University of Marketing &
Distribution Sciences
Gakuen-nishi-machi, Nishi-ku
Kobe 651-21, Japan