Mathematical Model for Preventive Maintenance Scheduling

G. Lorden
California Institute of Technology

D. S. Remer
Communications Systems Research Section and Harvey Mudd College

A model is formulated to describe the effect of the time interval chosen for preventive maintenance upon the frequency of failure and frequency of total maintenance (preventive and corrective). Trade-offs between these two frequencies are determined by computation of an optimal interval in the case where the failure distribution is known. For unknown distributions, an adaptive statistical technique is developed that converges to an optimal preventive maintenance interval. A numerical illustration is given.

I. Introduction

The proper goal of a preventive maintenance policy is to improve the availability and reliability of equipment. Such a policy is likely to be cost-effective, however, only if it is designed to take into account and to control the overall cost of corrective and preventive maintenance. DSN experience shows that preventive maintenance is a substantial component of cost (Ref. 1). This article is intended as a first step in the development of effective methods for reducing preventive maintenance costs by performing such maintenance only where and when it is most effective.

The basis for the method studied in this article is a probabilistic model of successive maintenance cycles. Each cycle ends with the performance of preventive maintenance at a scheduled time, or, if failure occurs earlier, with corrective maintenance. Whichever way a cycle ends, the successive cycles are assumed to be generated by independent trials of some failure distribution. The broad class of failure distributions considered allows for the possibility of higher failure rate early in the cycle, e.g., occasional bad effects resulting from maintenance activity.

After formulation of the model and derivation of necessary formulas in Section II, it is determined in Theorem 2 of Section III how to make optimal tradeoffs between the frequency of failure and frequency of maintenance (preventive and corrective combined) by choosing the right time interval for preventive maintenance based on a known failure distribution. Section IV considers a statistical technique for the case of unknown failure distribution. This technique modifies the choice of intervals for preventive maintenance as experience accumulates. The sequence of choices converges in probability to the optimum, as shown in Theorem 3. A numerical example illustrating the required computations is given in Section V. Additional remarks, including possibilities for further investigation, comprise Section VI.

II. The Model

Consider a fixed type of reparable equipment whose times between failures $T_1, T_2, \cdots$ are independent with distribution function F. Assume that F has a continuous density f, and let h be the failure rate function, $f/\bar{F}$, where $\bar{F} = 1 - F$. Thus, $h(t)$ represents the rate of failures in the interval between t and $t + dt$. Suppose that a time interval $m > 0$ is chosen for preventive maintenance. Then if a time m has elapsed since the last failure, preventive maintenance is performed. It will be assumed throughout that preventive maintenance, as well as corrective maintenance, restores the equipment to the condition where its time to next failure has distribution function F.

Theorem 1:

If $m > 0$ is the preventive maintenance interval, then

$$\text{frequency of failure} = \frac{F(m)}{\int_0^m \bar{F}}$$

and

$$\text{frequency of maintenance} = \frac{1}{\int_0^m \bar{F}}$$

the latter including both preventive and corrective maintenance.

Proof:

If $T_1, T_2, \cdots$ are the successive times between failures in the absence of preventive maintenance, then $X_1 = \min(m, T_1)$, $X_2 = \min(m, T_2), \cdots$ are the times between maintenances, and they are independent and identically distributed. Let N be the number of X's until the first failure, i.e., the first time $X_n = T_n$. Since N is distributed like the number of independent flips required to get the first "Heads," where $P(\text{Heads}) = P(T \leq m) = F(m)$, $EN = 1/F(m)$. Let $S_n = X_1 + \cdots + X_n$, $n \geq 1$, and note that the time of first failure is $S_N = X_1 + \cdots + X_N$, after which the whole process repeats itself. Since $ES_N = EN \cdot EX_1$ by Wald's equation for randomly stopped sums, the frequency of failure is

$$\frac{1}{ES_N} = \frac{1}{EN \cdot EX_1} = \frac{F(m)}{EX_1} = \frac{F(m)}{\int_0^m P(X_1 > t)\,dt} = \frac{F(m)}{\int_0^m \bar{F}}$$

Similarly, the frequency of maintenance is $1/EX_1$, which is given by the formula stated in the theorem, and the proof is complete.

III. Optimal Preventive Maintenance for Known F

Suppose that it is desired to choose a maintenance interval $m > 0$ to minimize

$$R(m) = \text{Frequency of failures} + c \cdot \text{frequency of maintenances}$$

$$= \frac{F(m) + c}{\int_0^m \bar{F}}$$

where $c \geq 0$ is chosen in advance. The choice of c determines the tradeoff between the two frequencies. If the choices made for different types of equipment are proportional to the relative costs of maintaining them, then the total spent on maintenance is distributed optimally – that is, minimizes the total failure rate of all equipment types. This is analogous to the determination of optimal allocations of spares to different types of equipment (Ref. 2).

This section considers the case where the failure distribution, F, is known and, hence, also f and h. Though not very practical, it is an instructive case to consider, and the results obtained provide a foundation for the more realistic formulation in the next section.

Theorem 2:

Assume that the failure rate function, h, is continuous and positive on $(0, \infty)$ and is "peak-free," i.e., is either monotonic or else, for some $d > 0$, is nonincreasing on $(0, d)$ and nondecreasing on $[d, \infty)$. Then the limit of h at $+\infty$, $h(\infty)$, exists (possibly infinite) and

(1) If $c \geq h(\infty)\, ET - 1$, then $R(m)$ is nonincreasing and bounded below by its limit at $+\infty$, $(1 + c)/ET$, which is attainable by choosing $m = +\infty$, whereas

(2) If $c < h(\infty)\, ET - 1$, then the solutions of $R(m) = h(m)$ are an interval $[m_1, m_2]$ (possibly $m_1 = m_2$) such that $R(m)$ is decreasing on $(0, m_1)$, constant on $[m_1, m_2]$, and increasing on $[m_2, \infty)$.

Proof:

The limit $h(\infty)$ exists since h is either nonincreasing or else is eventually nondecreasing. By routine calculation,

$$R'(m) = \frac{Q(m)}{\bar{F}(m)\left(\int_0^m \bar{F}\right)^2} \tag{1}$$

where

$$Q(m) = h(m)\int_0^a \bar{F} - [F(m)+c] \tag{2}$$

and

$$Q(b) - Q(a) = [h(b) - h(a)]\int_0^a \bar{F} + \int_a^b [h(b) - h]\,\bar{F} \tag{3}$$

It is clear from Eq. (3) that $h(t) \geq h(b)$ for all $t < b$ implies that $Q(t) \geq Q(b)$ also. Hence, Q is nonincreasing on any interval $(0, d)$ where h is. Also, if $h(b) \geq h(a)$, then $Q(b) \geq Q(a)$ and, hence, Q is nondecreasing wherever h is. Thus, Q is peak-free like h and by (2) is also continuous. Since $R(m) \to +\infty$ as $m \downarrow 0$, $R'(m) < 0$ for arbitrarily small positive m and by Eq. (1), therefore, $Q(m) < 0$ for arbitrarily small positive m. Since a peak-free function cannot assume a sequence of values negative-to-positive-to-nonpositive, evidently either Q is never positive or else it is negative on some interval $(0, m_1)$, zero on $[m_1, m_2]$, and positive (as well as nondecreasing) on (m_2, ∞). These two cases occur respectively, as $Q(\infty) = h(\infty)ET - (1+c)$ is ≤ 0 or > 0. Since, by (1), R' and Q have the same sign, the conclusions about $R(m)$ in the two cases follow immediately and the proof is complete upon noting that, by Eq. (2), $Q(m) = 0$ is equivalent to $R(m) = h(m)$.

Note that if h is differentiable, then

$$Q'(m) = h'(m)\int_0^m \bar{F}.$$

and Newton's method can be applied to solve $Q(m) = 0$ numerically.

IV. An Approach to the Optimal *m* When *F* Is Unknown

Under the assumptions of the preceding sections, we will now show how to choose successive maintenance intervals M_1, $M_2, \cdots$, based on accumulating experience, so that $\{M_n\}$ converges to the optimal interval $[m_1, m_2]$ in probability. The choice of M_{n+1} is based on the observations $Y_1, Y_2, \cdots, Y_n$, where $Y_i = \min(M_i, T_i)$ are the successive times between maintenances, and the choice also takes into account which of the Y_i's are failures (i.e., $T_i \leq M_i$). The method is based upon estimating the "survival function," $\bar{F}(x)$, at every stage by the Kaplan-Meier (Ref. 3) Product Limit Estimate (PLE)

$$K(x) = \prod_{t_k < x} \frac{\text{no. of observations surviving } t_k}{\text{no. of observations reaching } t_k}$$

where the t_k's are the points at which failures have been observed. Here the "number of observations reaching t_k" is just the number of Y's $\geq t_k$, while the "number of observations surviving t_k" is the same number minus the number of observed failures at t_k. It is convenient to modify the definition of $K(x)$ by stipulating that $K(x) = 0$ for $x >$ largest observation (failure or not). Note that $K(0) = 1$ and K is a step function with downward jumps at the points where failures have been observed, falling to 0 after the largest observation. It is well-known that $K(x)$ is a consistent estimator of $\bar{F}(x)$ (Ref. 1), i.e., if $K_n(x)$ denotes the estimate after n cycles, $Y_1, \cdots, Y_n$, then

$$K_n(x) \to \bar{F}(x) \text{ with probability one as } n \to \infty$$

for x such that $M_n > x$ for infinitely many n. In fact, this convergence is uniform for such x because the functions $K_n(x)$ are bounded and nonincreasing.

A recipe for choosing $M_1, M_2, \cdots$ can be given as follows. Let $R_n(\cdot)$ denote the estimate of the function $R(\cdot)$ obtained by using $K_n(x)$ in place of $\bar{F}(x)$ in the definition of R (note that $1 - K_n$ replaces F). Then let $M_1 = +\infty$ and for $n \geq 1$,

(1) Let M_n^* denote the value of m minimizing $R_n(m)$.

(2) Choose $M_{n+1} = M_n^*$ with probability $1 - 1/n$

$= +\infty$ with probability $1/n$.

Step (1) is computationally feasible since the numerator of $R_n(m)$ is constant for m between successive failure times, t_k, whereas the denominator increases, and it is easy to verify that the minimum of $R_n(m)$ is attained either at one of the failure times, t_k, or at the largest observation (failure or not). (Note also that the integral in the denominator of $R_n(m)$ is easily calculated since the integrand is K_n, a step function.) The randomization device used in step (2) provides a means of increasing M_n from time to time to gain information about possible m's larger than the ones used so far.

For the scheme of choosing $\{M_n\}$ just described, we obtain the following result:

Theorem 3:

If $b < m_1$, $a > m_2$, then $\lim_{n\to\infty} P(M_n < b) = 0$, $\lim_{n\to\infty} P(M_n > a) = 0$, and $\lim_{n\to\infty} P[R(M_n) > R(m_1) + \epsilon] = 0$ for $\epsilon > 0$.

Proof:

Since $P(M_n^* \neq M_n) = 1/n \to 0$, it suffices to prove all three conclusions with M_n^* in place of M_n. To prove the first conclusion for M_n^*, it suffices to show that for $b < m_1$, with probability one, all but finitely many M_n^*'s are $\geq b$.

Now, since

$$\sum_n P(M_n^* \neq M_n)$$

diverges, infinitely many M_n's equal $+\infty$, and, hence, $R_n(m) \to R(m)$ uniformly on $(0, b + 1]$.

Also, for all $m \leq b$, $R(m) \geq R(b) > R(\min(b + 1, m_1))$, so that, for sufficiently large n, by the uniform convergence,

$$R_n(m) > R_n(\min(b + 1, m_1)) \text{ for all } m \leq b$$

which implies that $M_n^* > b$. Thus, only finitely many M_n^* are less than b, and the first limit in Theorem 3 is proved.

The second limit is shown to be zero by a similar argument. The third limit is zero since by Theorem 2 $R(m) > R(m_1) + \epsilon$ only if $m < b$ or $m > a$ for some $b < m_1$ or $a > m_2$.

V. A Numerical Example

To illustrate the computations needed for the method of the preceding section, suppose that successive cycles of lengths (in days) 47, 26, 26, 19, 27, 16, 18, 20, 20, 35 are observed, where the underlinings denote failures. Assume that $c = 0.5$. The necessary computations are shown in Table 1, and K and the estimated R are graphed in Figs. 1 and 2. Note that the computations only need to be performed at time points where failures occurred and at the largest observation (failure or not). The minimum $R(t)$ is at the largest observation, 47. Thus, $M_{10}^* = 47$ and this value would be chosen as the preventive maintenance time for the next cycle, unless the randomization produced $M_{11} = +\infty$ (the probability of this being 0.1), in which case there would be no preventive maintenance in the next cycle, and the cycle would end at the next failure. At the end of the next cycle, new computations of the entries in Table 1 would be required. If the cycle ended at 17, for example, the "16" column would be unchanged, but the "19" and "26" columns would be recomputed. Also, if the cycle ended with a failure (at a new t-value), then a new column would be inserted for that t-value.

This updating of the columns after each cycle is not difficult because $K(t)$ is obtained by multiplying the value in the preceding column by a fraction, and the integral of $K(x)$ up to t is obtainable by adding the area of a rectangle to the integral in the preceding column.

VI. Additional Remarks

In practice, the situation is usually slightly more complicated than that described in the preceding section, because one has several pieces of the same type of equipment and simultaneously must set M_n's and accumulate experience from all of them. It is not hard to modify the recipe, however, to deal with this situation. One can simply recalculate $K(\cdot)$ after each observation (on any of the pieces) and calculate the next M_n desired. If one or more pieces have already exceeded an elapsed time of M_n since their last maintenance, then perform preventive maintenance on them. Thus, one sometimes observes longer cycles than the recipe would call for, but there is no significant change needed in the proof of Theorem 3 or in the carrying out of the recipe.

It is interesting and perhaps useful to try to relax the assumption that preventive maintenance restores the equipment to its original failure distribution, F – say, to allow a separate contribution to the failure rate depending upon the age of the equipment. It is straightforward to modify the "known F" analysis of Section 2 to accommodate this sort of extension – even if the age-contribution is unknown (since it is unaffected by the choice of m and, hence, acts merely as a sort of "background radiation" of failures). The extension in the "unknown F" case, however, seems more difficult. It is perhaps helpful to assume that the age-dependent failure rate is known.

Another promising approach to the determination of preventive maintenance strategies is the use of measurements of "indicator variables" reflecting the need for preventive maintenance. These variables might be levels of contamination, pressure, vibration, etc., or various measures of performance like the rate of random errors. By measuring such variables, one can expect to anticipate failures that could be prevented (or, at least, postponed) by timely maintenance.

References

1. Remer, D. S., and Lorden, G., "Initial Economic and Operations Data Base for DSS-13 Automation Test, *DSN Progress Report 42-49*, Jet Propulsion Laboratory, Pasadena, Calif.

2. I. Eisenberger and Lorden, G., "Dynamic Spares Provisioning for the DSN," *DSN Progress Report 42-42*, Jet Propulsion Laboratory, Pasadena, Calif.

3. Kaplan, E. L., and Meier, P., "Nonparametric Estimation from Incomplete Observations," *J. Amer. Statist. Assoc.*, 53, p. 457-481, 1958.

Table 1. Computational results for illustration

Value of t	16	19	26	47
No. reaching t	10	8	5	1
No. surviving t	9	7	4	1
$K(t)$	1	0.9	0.788	0.63
$\int_0^t K(x)\,dx$	16	18.7	24.21	37.44
estimated $R(t)$	0.0313	0.0321	0.0294	0.0232

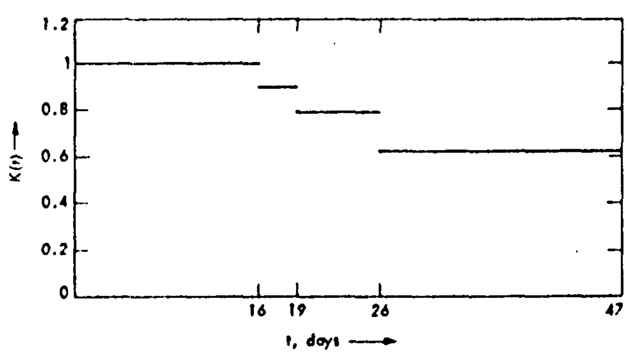

Fig. 1. $K(t)$

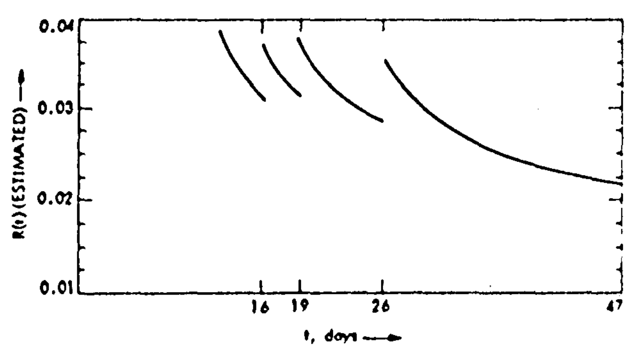

Fig. 2. Estimated $R(t)$