

OPTIMAL MONITORING STRATEGIES FOR SLOWLY
DETERIORATING REPAIRABLE SYSTEMS

Peter Kubat
GTE Laboratories Incorporated
40 Sylvan Road
Waltham, MA 02254

C.Y. Teresa Lam
Department of Industrial
and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109

Technical Report 89-27

September 1989

Optimal Monitoring Strategies for Slowly Deteriorating Repairable Systems

Peter Kubat
GTE Laboratories Incorporated
40 Sylvan Road
Waltham, MA 02254

C. Y. Teresa Lam
Department of Industrial
and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109

Abstract

Many customer reported troubles can be prevented if the deterioration of network components is recognized early and corrective actions are taken before potential troubles are observed by customers.

In this note we present a simple model for determination of an optimal "action limit" in slowly deteriorating repairable system. The performance of such a system is assumed to be characterized by a single parameter which is continuously being monitored. The underlying deterioration process is assumed to be governed by a Brownian motion process with a positive drift. When the measured value of the parameter reaches the so called "action limit," the repair/replacement procedure is initiated. The optimal action limit is derived so that the expected long run average total cost is minimized. Some simple numerical examples illustrate the model and the optimization.

1 Introduction

In order to provide a high quality and reliable service, many components of today's complex systems are continuously monitored to determine if they function properly and to detect malfunction as soon as possible. For example, in telecommunication networks and/or computer systems many critical components are constantly being monitored in order to detect any performance degradations which may later result in a component or system failure and consequently in interruption of service to the customer. When a faulty condition is detected, evasive actions can be taken (as, for instance, switching to spare component) to correct the fault well before potential troubles will be experienced by customers. This research was motivated by the need to develop effective monitoring strategies to guarantee trouble-free service of a communication/computer network and enhance its degree of fault tolerance.

Not all the failures can, of course, be predicted by the continuous monitoring. Failures caused by random phenomena, such as stochastic failure of electronic components, storms, accidents, malicious acts, careless construction crews and other exogenous sources are typical examples. On the other hand, many failures may be preventable. For instance, failures of equipment in

which some parts deteriorate slowly over time (such as copper wire pairs or terminal connections corroding in moist environment) may be prevented if the problem is recognized early and corrected in time.

In order to detect any possible performance degradations of network components as soon as possible some relevant parameters of the component will be either continuously or frequently monitored to make sure that the parameter values are within certain well specified acceptable limits. When the monitor will detect a significant change in the monitored parameters, network control will be notified and a sequence of corrective actions (such as standby components activated, traffic rerouted, repair procedure initiated, etc.) will then be undertaken.

This paper develops a methodology for determination of optimal monitoring strategies for a repairable system whose performance is slowly deteriorating over time. The following monitoring and repair setting will be assumed: The performance of the system at the time t is characterized by a single parameter $X(t)$. This parameter is continuously monitored. If $X(t)$ is bigger than a certain threshold, say U , customers notice the degradation and complain. Since $X(t)$ is continuously monitored, some level of degradation can be observed before the customers notice. If the degradation is caught early, then frequently the system can be fixed before $X(t)$ reaches U and thus the majority of customer complaints can be eliminated.

Without loss of generality, assume that the smaller the value of $X(t)$, the better is the system performance, in other words, the system performance is a nonincreasing function of $X(t)$. Examples of $X(t)$ include BER (Bit Error Rate), frequency of ARQ (Automatic Repeat Request) or CRC (Cyclic Redundancy Code) checks. If, in some applications, such as when we assume that the quality of a telephone connection is measured by signal/noise ratio (the higher the S/N ratio, the better the quality) we take $X(t)$ to be a negative of the S/N ratio measurement.

The system is then judged to have a satisfactory performance if the value of $X(t)$ lies below the threshold U . We will assume that the deterioration process $X(t)$ can be modeled by a Wiener (Brownian) process with a positive drift.

To correct the degradation, (hopefully before it is observed by a customer) we consider the following repair scenario: When, for the first time, the process $X(t)$ reaches a certain predetermined level, called "action limit" (lower than the threshold), an action "issue a trouble ticket," is taken and a repair (overhaul, replacement, etc.) procedure is initiated. Clearly, there is a cost involved with each repair. The repair cost includes the actual repair cost, cost of spare parts, administration and dispatching cost, etc. The time to repair completion is assumed to be a random variable and the service provider will be penalized if the repair is not completed before the process reaches the threshold. We derive the optimal action limit which will minimize the expected long run average total cost per unit time.

In the recent years, new approaches for maintenance policies for the control of production systems which are subject to deterioration over time have been considered by Lee and Rosenblatt, (1988), Paté-Cornell, Lee and Targas (1987) among many others. Models for continuous monitoring of production control with warning signals was considered by Lee, Moinzadeh and Targas (1986). These three papers contain essentially a complete bibliography on the subject.

Rapid advances in microprocessor technology and data communication networks made effective and instant monitoring possible for many components of complex and geographically distributed systems. Today's network control centers monitor thousands of subsystems and components, yet, even the traditional statistical quality control routines to detect shifts from "in control" to "out-of-control" state are not widely implemented.

In a related problem, Eu and Rollins (1988) consider a sampling strategy to estimate real-time error performance of digital links and suggest determination of alarm conditions for the purpose of trouble prevention. Models for effective scheduling of periodic diagnostic tests (monitoring from time to time) in communications systems were considered by Kubat (1988) and by Rubin and Zhang (1988).

Models of monitoring and controlling of production processes, based on Brownian motion were considered by Antelman and Savage (1965) and later also discussed by Ross (1983) under the assumption that the repair process is instantaneous. The monitoring models considered in this paper differ from the Antelman and Savage models in both the monitoring strategy and in the cost function. Lee, Moinzadeh and Targas (1986), consider a deterministic deterioration function which is zero until a fault develops at a random time.

Although this research was mostly motivated by the monitoring needs in communication systems, it can be useful in other fields as well. Some examples are: production processes in which cutting tools are periodically replaced because of the tools wear and tear, jet engines, materials changing chemical structure over time, lasers slowly losing power over time, electric and electronic systems depending on replaceable power source, close pressure systems, etc.

2 Continuous Monitoring Model

Consider a Brownian motion $\{X(t), t \geq 0\}$ with positive drift μ and variance σ^2 , starting at the origin, i.e., $X(0) = 0$. Without loss of generality, we set the threshold $U = 1$, and denote by p ($0 \leq p \leq 1$) the action limit. Let $T_{a,b}$ denote the first passage times for the process to reach level b starting from level a , ($a < b$). Specifically, we denote by T_p and T_1 the first passage times for the process to reach p , and 1 respectively starting from the origin (i.e., $T_p = T_{0,p}$ and $T_1 = T_{0,1}$); whereas $T_{p,1}$ denotes the first passage time to reach 1 starting from level p ; $G(t)$ denotes the distribution of $T_{p,1}$. Let $E[T_p]$ be the expected first passage time from zero to p (see Figure 1 for details). When the process $X(t)$ reaches level p the repair process is initiated. The trouble ticket is issued, additional tests are made and finally the repairman is notified and dispatched. Let R be the time from the instant in which the action limit is hit and trouble ticket issued until the repair is successfully completed. This time, sometimes called a "lead time" is a random variable and has distribution $F(r)$. Note that the lead time will include all the queueing, administration, system testing, dispatching, transportation and other delays which may occur during the repair cycle. We assume that the repair time itself is small and will be included in R . Furthermore, during the actual repair, the customer will be taken out of service only for a negligible amount of time and thus all the costs associated with this service interruption can be

neglected as well.

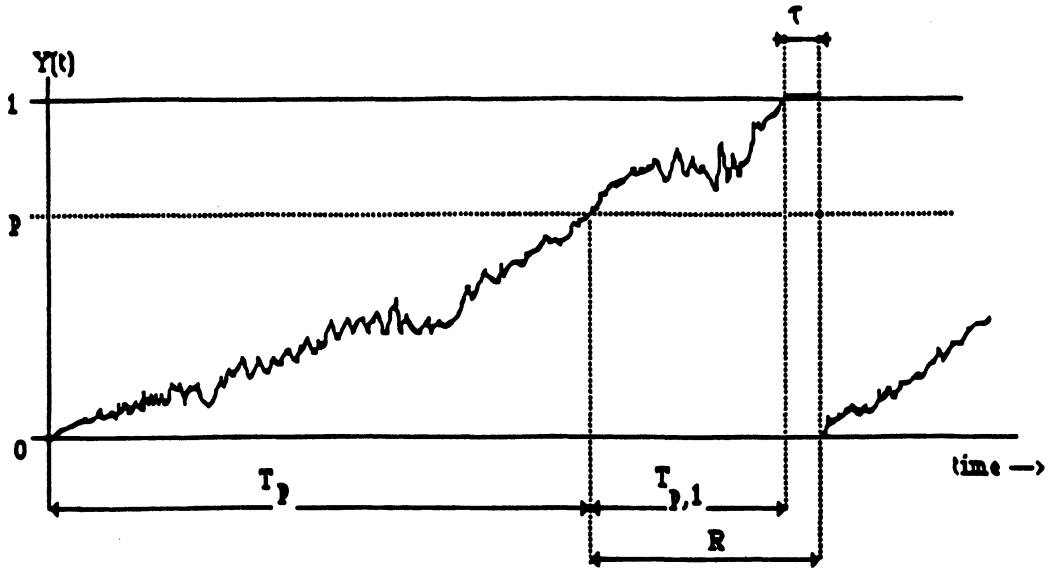


Figure 1. A typical realization of process $Y(t)$.

From Karlin and Taylor (1975, p. 362), we get that the Laplace transform of $T_{a,b}$

$$(2.1) \quad E[\exp(-\theta T_{a,b})] = \exp\left(-\frac{(b-a)}{\sigma^2} \left(\sqrt{\mu^2 + 2\sigma^2\theta} - \mu\right)\right)$$

where θ denotes the argument of the Laplace function. Consequently

$$(2.2) \quad E[T_p] = -\frac{d}{d\theta} E[\exp(-\theta T_p)] |_{\theta=0} = \frac{p}{\mu}$$

Define

$$Y(t) = \begin{cases} X(t) & \text{if } t < T_1 \\ 1 & \text{if } t \geq T_1 \end{cases}$$

and

$$\tau = \begin{cases} R - T_{p,1} & \text{if } R \geq T_{p,1} \\ 0 & \text{if } R < T_{p,1} \end{cases}$$

The process $Y(t)$ is called the truncated Brownian motion and τ can be interpreted as the time that the truncated process $Y(t)$ spends at the state 1, or equivalently, the time the system will be in the unacceptable state waiting for the repair completion. The times at which the repair is just completed, can be considered as regeneration points of the underlying regenerative renewal process. The time interval between two successive regeneration points is called the cycle time.

Suppose that every time a repair is initiated, a fixed cost c_1 is incurred. This cost may be derived as an average cost of a trouble ticket disposition. During the period when the system is in the unacceptable state (i.e., $Y(t) = 1$) additional cost is accumulated. This cost is linearly proportional to the length of that period and c_2 denotes the per unit time cost when the system

is in unacceptable state. This cost may represent the cost of forgone revenues, opportunity loss, cost of customer inconvenience, etc.

After the repair completion, the process $Y(t)$ is put back at the origin and a new cycle starts. Clearly,

$$(2.3) \quad E[\text{length of a cycle}] = E[T_p] + E[R]$$

and

$$(2.4) \quad E[\text{cost of a cycle}] = c_1 + c_2 E[\tau],$$

where

$$(2.5) \quad \begin{aligned} E[\tau] &= \int_{0+}^{\infty} P(\tau > a) da \\ &= \int_{0+}^{\infty} P(R - T_{p,1} > a) da \\ &= \int_{0+}^{\infty} \int_0^{\infty} P(R > a + t) dG(t) da \\ &= \int_{0+}^{\infty} \int_0^{\infty} \int_{a+t}^{\infty} dF(r) dG(t) da \end{aligned}$$

Following the theory of regenerative processes with rewards (Ross (1983), pp. 81-83), we see that the expected long run average cost per unit time is

$$(2.5) \quad \gamma(p) = \frac{c_1 + c_2 E[\tau]}{\frac{p}{\mu} + E[R]}$$

In the case that R is exponentially distributed with parameter λ , $E[\tau]$ is greatly simplified. Here we have:

$$(2.7) \quad \begin{aligned} E[\tau] &= \int_0^{\infty} \int_0^{\infty} \int_{a+t}^{\infty} \lambda \exp(-\lambda r) dr dG(t) da \\ &= \int_0^{\infty} \exp(-\lambda a) \left[\int_0^{\infty} \exp(\lambda t) dG(t) \right] da \\ &= \int_0^{\infty} \exp(-\lambda a) E[\exp(\lambda T_{p,1})] da \\ &= \frac{1}{\lambda} E[\exp(-\lambda T_{p,1})] \\ &= \frac{1}{\lambda} \exp \left\{ -(1-p) \frac{\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu}{\sigma^2} \right\} \end{aligned}$$

so that

$$\begin{aligned}
 \gamma(p) &= \frac{c_1 + c_2 E[\tau]}{\frac{p}{\mu} + \frac{1}{\lambda}} \\
 (2.8) \quad &= \frac{c_1 \mu \lambda + c_2 \mu \exp \left\{ -(1-p) \frac{\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu}{\sigma^2} \right\}}{p\lambda + \mu}
 \end{aligned}$$

Our purpose now is to choose an optimal p so as to minimize the expected long run average cost per unit time (2.8) and we must have that $0 \leq p \leq 1$.

Differentiating $\gamma(p)$ with respect to p and putting $\gamma'(p) = 0$, we obtain the equation

$$\begin{aligned}
 (2.9) \quad &c_2 \lambda \left(\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu \right) p + c_2 \left(\mu \sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu^2 - \sigma^2 \lambda \right) \\
 &= \sigma^2 \lambda^2 c_1 \exp \left\{ (1-p) \frac{\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu}{\sigma^2} \right\}
 \end{aligned}$$

First note that

$$(2.10) \quad c_2 \lambda \left(\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu \right) > 0$$

and

$$\begin{aligned}
 (2.11) \quad &(\sigma^2 \lambda + \mu^2)^2 = \sigma^4 \lambda^2 + \left(\mu \sqrt{\mu^2 + 2\sigma^2 \lambda} \right)^2 \\
 &> \left(\mu \sqrt{\mu^2 + 2\sigma^2 \lambda} \right)^2
 \end{aligned}$$

implies that

$$(2.12) \quad c_2 \left(\mu \sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu^2 - \sigma^2 \lambda \right) < 0$$

i.e. the left hand side of equation (2.9) is a straight line with positive slope and negative intercept, whereas the right hand side of (2.9) is an exponential function decreasing as p increases and it intercepts the positive y -axis. Thus the straight line and the exponential function must intersect at a point, this gives us the solution of equation (2.9) and it must be greater than 0 (see Figure 2). This solution is less than or equal to 1 if and only if the intersection happens at p less than or equal to 1, this is equivalent to the following constraint on the parameters:

$$(2.13) \quad c_2 \lambda \left(\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu \right) + c_2 \left(\mu \sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu^2 - \sigma^2 \lambda \right) \geq \sigma^2 \lambda^2 c_1$$

It remains to show that the solution is indeed the minimum of our objective function, this can be done by showing that our function is convex. Differentiating $\gamma'(p)$ with respect to p again to obtain

$$\begin{aligned}
 \gamma''(p) &= \frac{1}{\sigma^4 (p\lambda + \mu)^3} \left\{ c_2 \mu A \left[(p\lambda + \mu) \left(\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu \right) - \lambda \sigma^2 \right]^2 + c_2 \mu A \lambda^2 \sigma^2 + 2c_1 \mu \lambda^3 \right\} \\
 &> 0
 \end{aligned}$$

where

$$A = \exp \left\{ -(1-p) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\}$$

The convexity of the objective function further implies that if constraint (2.13) is not satisfied, then the objective function is minimized at the point $p = 1$.

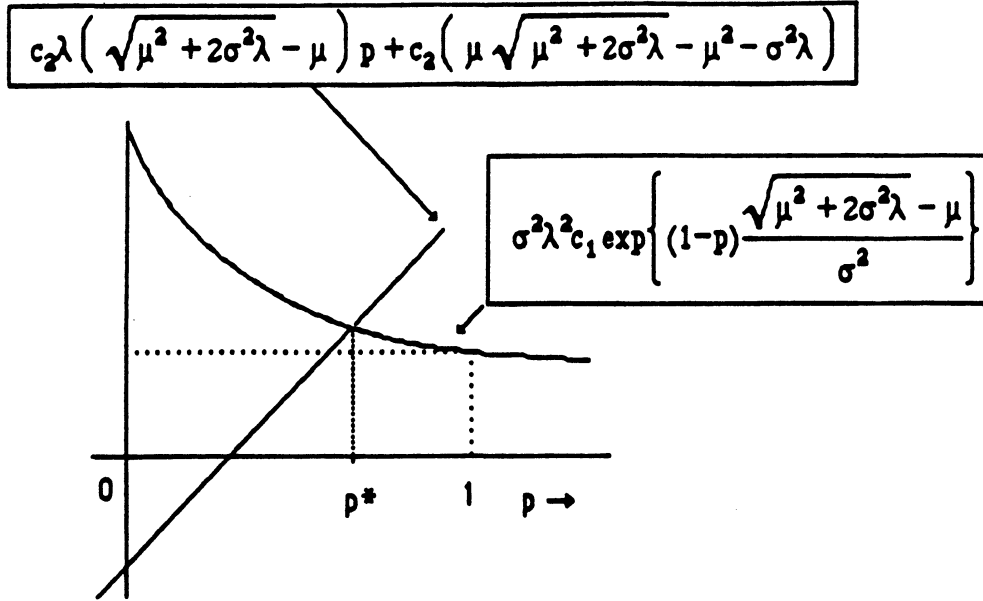


Figure 2. Solution of the equation (2.9).

3 Sensitivity Analysis

It is of interest to study not only the change of the expected long run average cost per unit time as the action limit p changes, it is also important to see how the other parameters affect the long run average cost per unit time. For instance, λ is the rate at which the repairman responds, it is therefore reasonable to expect that the long term cost increases as λ decreases. Direct differentiation shows that in the case when $c_1 > 0$, γ does increase as λ decreases, however in the case when $c_1 = 0$, the behavior of γ as λ changes is much more complicated. When we set the action limit $p = 1$, i.e. the repairman is only called at the moment when the system is in the unacceptable state, then γ increases as λ decreases if and only if the ratio of the per unit time cost for the system to be in the unacceptable state to the fixed cost for the repairman to come is greater than μ , the deteriorating percentage rate of the system. Strict forward differentiations also shows that, as expected, γ increases as c_1, c_2, μ or σ increase.

Special case when $c_1 = 0$. When $c_1 = 0$, i.e. the fixed cost of the repair is equal to 0, the right hand side of equation (1) is identically equal to 0 and the optimal p can be solved as an

explicit function of the other parameters. In this case, the solution p^* is

$$(3.1) \quad p^* = \frac{\mu^2 + \sigma^2 \lambda - \mu \sqrt{\mu^2 + 2\sigma^2 \lambda}}{\lambda (\sqrt{\mu^2 + 2\sigma^2 \lambda} - \mu)}$$

Note that p^* is independent of c_2 , this is because when $c_1 = 0, c_2$ becomes a scaling factor for the objective function. In this case, p^* lies in the range $[0,1]$ if and only if $\sigma^2 - 2\mu \leq 2\lambda$. Direct differentiations of p^* with respect to σ, μ or λ allow us to find out how we should set our action limit p^* as other parameters change. As expected, p^* decreases as either λ or μ increase, i.e. when either the deteriorating percentage rate of the system or the rate of arrival of the repairman increase, we should set a lower action limit so as to minimize the long term average cost per unit time. However, as the diffusion coefficient σ^2 increases, p^* increases. This means that when the variability of the deteriorating percentage of the system increases, a higher action limit is required.

4 Numerical Examples

Example 1. Consider $\mu = 0.01$, this means that on average it takes 100 days before the system reaches the threshold and let us choose $\sigma = 0.05$. We take $\lambda = 1, 0.5$ and 0.25 , i.e., the mean time for the repairman to come ranges from 1 day to 4 days. We select the repair cost $c_1 = 100(\text{\$})$, a typical cost of average repair call (with overhead) in the outside plant, and select $c_2 = 2,000$ ($\text{\$/day}$).

The optimal action limit p^* is given for the above mentioned values of λ are:

λ	p^*
1	0.75
0.5	0.65
0.25	0.45

The corresponding long run expected cost per day is plotted as a function of p and is shown in Figure 3.

Example 2. Here we select $\lambda = .5$ and we will investigate the effect of the failure rate μ on the long run average cost. We take $\mu = 0.001, 0.005$ and 0.01 , i.e., the mean time to failure ranges from 100 days to 1,000 days. As in the previous example we set $\sigma = 0.05, c_1 = 100(\text{\$})$ and $c_2 = 2,000$ ($\text{\$/day}$). The optimal action limit p^* is given for the above mentioned values of μ are:

μ	p^*
0.01	0.65
0.005	0.65
0.001	0.75

The corresponding long run expected cost per day is plotted as a function of p and is shown in Figure 4.

From Figure 3 and 4 it is interesting to observe that as λ or μ increases, p^* increases. This is reasonable because as the failure rate or lead time increases, the action to repair has to be taken sooner to protect against possible outages. Furthermore, the cost function per unit time γ is rather flat in the neighborhood of the optimal point p^* , which indicates that the optimal action limit is robust.

5 Behavior of p^* as a function of c_1 and c_2

In examples 1 and 2, we studied how we should set our action limit in terms of the parameters λ, μ or σ , but keeping c_1 and c_2 fixed. In this section, the implicit equation (2.9) of p is used to find out the changes in our action limit for different values of c_1 and c_2 .

First note, that equation (2.9) can be rewritten as

$$(5.1) \quad \begin{aligned} & \frac{c_2}{c_1} \left\{ \lambda \left(\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu \right) p + \left(\mu \sqrt{\mu^2 + 2\sigma^2\lambda} - \mu^2 - \sigma^2\lambda \right) \right\} \\ & = \sigma^2 \lambda^2 \exp \left\{ (1-p) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\} \end{aligned}$$

where

$$(5.2) \quad \mu \sqrt{\mu^2 + 2\sigma^2\lambda} - \mu^2 - \sigma^2\lambda < 0$$

for all choices of μ, λ and σ .

The left hand side of equation (5.1) is again a straight line with positive slope and negative intercept and the right hand side is a decreasing exponential function intersecting the positive y -axis, the intersection of these two curves gives the optimal action limit for our problem. Note that as c_2/c_1 increases, the straight line rotates in a counterclockwise direction about a fixed point and the optimal p becomes smaller. On the other hand, as c_2/c_1 decreases, the straight line rotates in a clockwise direction about the same fixed point, the x -intercept of the straight line, so that the optimal action limit increases. It is interesting to note that if the ratio c_2/c_1 is a constant, then the optimal p is independent of c_1 or c_2 . However, if we hold c_1 fixed, then the action limit decreases as c_2 increases. This is reasonable because when the per unit time cost for the system to be in the unacceptable state increases and the other parameters are fixed, we would want to set a lower action limit so that the repairman can come earlier. If c_2 is now fixed and c_1 varies, optimal p should increase as c_1 increases, this is because when the fixed cost for the repairman increases, we do not want to call the repairman as often, in particular, we want the cycle time to be longer so as to minimize the long run average cost per unit time.

6 Generalization

(a) **Lead Time Cost.** So far we have assumed that there is a fixed cost c_1 everytime when the repairman is called, we were looking at

$$(6.1) \quad \min_p \left[\frac{c_1}{\frac{p}{\mu} + \frac{1}{\lambda}} + \frac{c_2}{\lambda} \frac{\exp \left\{ -(1-p) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\}}{\frac{p}{\mu} + \frac{1}{\lambda}} \right]$$

The second term in the expression above can be thought of as the long run average cost per unit time due to the system being in the unacceptable state and the first term is the long run average cost per unit time of the repairman. The sum then gives us the long run average cost per unit time of maintaining the system. The assumption that there is a fixed cost c_1 everytime the repairman is called can be generalized. Let $c(\lambda)$ be the per unit time cost of maintaining the cost parameter at level λ , our new minimization problem becomes

$$(6.2) \quad \min_{p,\lambda} \left[c(\lambda) + \frac{c_1}{\frac{p}{\mu} + \frac{1}{\lambda}} + \frac{c_2}{\lambda} \frac{\exp \left\{ -(1-p) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\}}{\frac{p}{\mu} + \frac{1}{\lambda}} \right]$$

It is reasonable to assume that $c(\lambda)$ is an increasing function of λ , i.e. the per unit time cost of maintaining the control parameter λ increases as the rate of repairs increases. If $c(\lambda)$ is assumed to be independent of the other parameters, then the minimization (6.2) is the same as

$$(6.3) \quad \min_{\lambda} \left[c(\lambda) + \frac{c_1}{\frac{p^*}{\mu} + \frac{1}{\lambda}} + \frac{c_2}{\lambda} \frac{\exp \left\{ -(1-p^*) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\}}{\frac{p^*}{\mu} + \frac{1}{\lambda}} \right]$$

when p^* is the optimal p in the minimization (6.1). Note here that p^* depends on λ in the implicit equation (2.9) and so (6.3) is not easy to differentiate.

In the case when $c_1 = 0$, we want to minimize the function $f(\lambda)$ where

$$(6.4) \quad f(\lambda) = c(\lambda) + \frac{c_2}{\lambda} \frac{\exp \left\{ -(1-p^*) \frac{\sqrt{\mu^2 + 2\sigma^2\lambda} - \mu}{\sigma^2} \right\}}{\frac{p^*}{\mu} + \frac{1}{\lambda}}$$

and p^* is an explicit function of λ .

(b) **Non-exponential Lead Time Distribution.** A close expression for $E[\tau]$, given by (2.7), can be also obtained if the distribution of R is a mixture of exponentials, i.e., the density of R

is given by

$$f_R(r) = \sum_{i=1}^n \xi_i \lambda_i \exp\{-\lambda_i r\}, \quad \sum_{i=1}^n \xi_i = 1, \quad \xi_i > 0.$$

Many distributions can be approximated by a mixture of exponentials and thus our problem is fairly general. In this case, however, the minimization of the expected long run average cost (2.8) is more complicated, but solvable using standard numerical techniques.

7 Concluding Remarks

In this paper we presented a simple continuous monitoring scheme for a repairable system. The underlying deterioration process of the system is assumed to be governed by a Brownian motion process with a positive drift. An optimal action limit has been derived and its sensitivity with respect to the system parameters and costs have been studied. The model and the derivation of the action limit have been illustrated on a few simple numerical examples.

In practice it may be impossible or too costly to continuously monitor the system performance, it would be useful to extend our model so that the relevant parameters of the systems are only monitored from time to time. Models for periodic monitoring and action limits are currently under development.

References

- Antelman, G. R., and I. R. Savage, (1965), "Surveillance Problems: Wiener Processes," *Naval Research Logistic Quarterly*, Vol. 12, 35-55.
- Eu, J. H., and W. W. Rollins, (1988), "A Sampling Measurement System for Continuous Real-Time Error Performance Monitoring of Live Traffic Digital Transmission Systems," Proceedings of IEEE Network Operations and Management Symposium (NOMS'88), New Orleans, LA.
- Karlin, S. and H. M. Taylor, (1975), *A First Course in Stochastic Processes*, Second Edition, Academic Press, New York.
- Kubat, P., (1988), "Optimal Level of Diagnostics in Distributed Communication Systems," *Europ. J. Oper. Res.*, Vol. 36, pp. 346-352.
- Lee, H. L., K. Moinszadeh, and G. Targas, (1986), "A Model for Continuous Production Control with Warning Signals to Fault Occurrences," *J. Opl. Res. Soc.*, Vol. 37, No.5, pp. 515-523.
- Lee, H. L. and M. J. Rosenblatt, (1988), "Economic Design and Control of Monitoring Mechanisms in Automated Production Systems," To appear in *IIE Transactions*.
- Paté-Cornell, M. E., H. L. Lee, and G. Targas, (1987), "Warnings of Malfunction: The Decision to Inspect and Maintain Production Processes on Schedule or on Demand," *Management Science*, Vol.33, No. 10, pp. 1277- 1290.

Ross, S.M. (1983), *Stochastic Processes*, Wiley, New York

Rubin, I. and Z. Zhang, (1988), "Test Scheduling for Communication and Queuing Processors,"
Proceedings of IEEE Network Operations and Management Symposium (NOMS'88), New Orleans, LA.

Figure 3: $\mu = 0.01, \sigma = 0.05, c_1 = 100, c_2 = 2000$

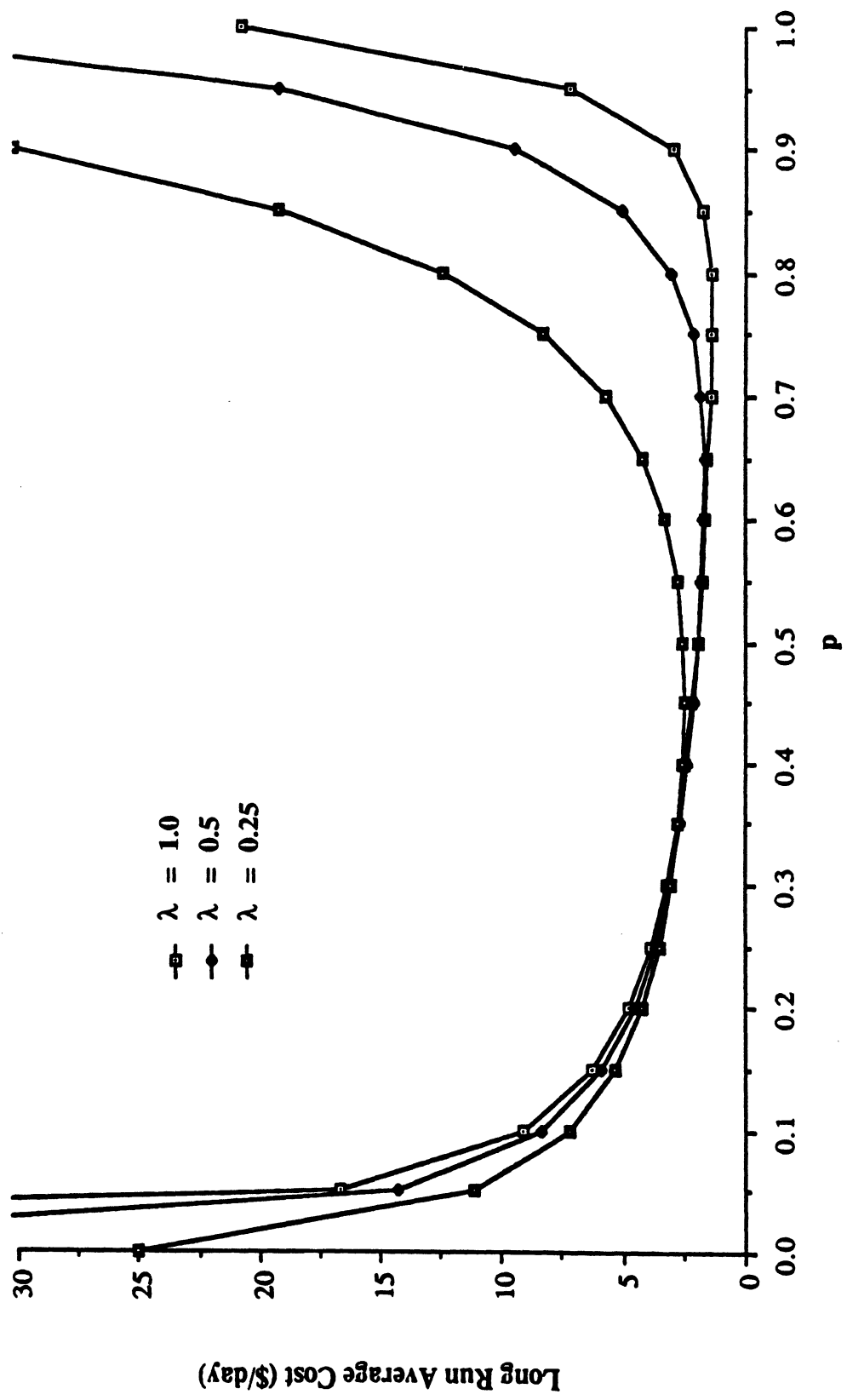


Figure 4: $\lambda = 0.5, \sigma = 0.05, c_1 = 100, c_2 = 2000$

