

## **General Disclaimer**

### **One or more of the Following Statements may affect this Document**

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

*Preprint*

DAA  
NAG 1-296

# OPTIMAL DESIGN AND USE OF RETRY IN FAULT TOLERANT REAL-TIME COMPUTER SYSTEMS<sup>1</sup>

Yann-Hang Lee and Kang G. Shin

## ABSTRACT



In this paper, we present a new method for (i) determining an optimal retry policy and (ii) using retry for fault characterization.

First, we derive an optimal retry policy for a given fault characteristic, which determines the maximum allowable retry durations so as to minimize the total task completion time. Then, we carry out the combined *fault characterization* and *retry decision*, in which the characteristics of fault are estimated simultaneously with the determination of the optimal retry policy. We have developed two solution approaches; one is based on the point estimation and the other on the Bayes sequential decision. The maximum likelihood estimators are used for the first approach, and the backward induction for testing hypotheses in the second approach.

We also present numerical examples in which all the durations associated with faults (i.e. active, benign, and inter-failure durations) have monotone hazard rate functions, e.g., exponential, Weibull and gamma distributions. These are standard distributions commonly used for modeling and analyses of faults.

Categories and Subject Descriptors: B.2.3 [Arithmetic and Logic Structures]: Reliability, Testing and Fault-Tolerance -- *hazard rate function, recovery overhead, optimal retry policy, fault characteristic*; G.3 [Probability and Statistics] -- *estimation, censored sampling, likelihood ratio, sequential or Bayes decision problem, hypotheses testing.*

RECEIVED  
MAY 17 11 11:47  
F. I. S. LIBRARY

<sup>1</sup>This work was supported in part by NASA under Grant NAG 1-296. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NASA.

Authors' address: Division of Computer Science and Engineering, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109. All correspondence should be addressed to Prof. Kang G. Shin at the above address.

## 1. INTRODUCTION

There are three types of fault in computer systems: *transient*, *intermittent*, and *permanent* [1]. Transient faults die within a certain time of their generation, intermittent faults cycle between being active and inactive, and permanent faults are (as the term indicates) permanent. It has been found that permanent faults form but a small fraction of the faults in computer systems [2,3]. This makes the purging of any faulty components as soon as they have been discovered an inefficient means for handling redundancy. If the active duration of a transient or intermittent fault is short, the continuation of the task with the same resource after the disappearance of the fault may be more efficient than that of using other recovery methods. Unfortunately, it is impossible to tell at its first occurrence whether or not the fault is permanent and also impossible to know its active duration if the fault is intermittent or transient. Moreover, it would be much more efficient (time-wise) and accurate to characterize faults on-line, and then take the appropriate recovery actions. In this paper, we propose (i) determining an optimal retry policy so as to minimize the task completion time, and (ii) using retry in conjunction with statistical estimation and decision theory to characterize faults. We obtain the optimal retry duration in the face of uncertainty about the nature of a detected fault. Since our focus is on real-time systems, we are principally concerned with skewing the density function of the task completion time as much to the left as possible. For this reason, we shall concentrate on maximizing reductions in response time.

As the term implies, retry consists of restoring the affected process to some fault-free initial state, and then re-running it on the same processor. Clearly, retry is only applicable when the error induced by a fault is confined and the process can be restored to integrity. The most efficient means for fault confinement are signal-level detection

mechanisms which can detect faults immediately upon their occurrence [4]. For the process to be restorable to integrity, scratchpad registers are needed. Results obtained by Carter, *et al.* [5] indicate that self-checking and retry mechanisms can be incorporated into processors inexpensively, and without substantially degrading performance.

Currently, several commercial machines incorporate retry. In the Honeywell 6000 [6], instruction retry is reported to approach an effectiveness rate of 100%. Retry in the IBM 360 and 370 series machines is widely used in the peripheral areas (I/O and storage) as well as in the central processor [7]. The UNIVAC 1100/60 uses a hardware timer that goes off after an interval judged to be long enough to allow transient faults to die out, upon which retry can be effected [8]. However, no discussion or justification about the retry duration or the number of retry attempts used has been addressed.

The usefulness of retry mechanisms arises, as we said above, from (i) the smallness of the proportion of permanent faults in any computer system, and (ii) the fast recovery from non-permanent faults and thus the small task completion time. In the case of a permanent fault, to retry a process on the affected processor is worse than useless: it is a waste of time. To hasten the completion of the executing task when a fault is detected, we must control the duration of retry to maximize the difference between the expected gain in response time that results from using retry when the fault is transient or intermittent (in some cases), and the expected loss that results from using it when the fault is permanent or intermittent (in other cases). Our object in this paper is to derive the maximum allowable retry duration  $r'$  when a fault is detected. If the retry succeeds with this duration, the execution continues. If not, other methods for error recovery, e.g., rollback or restart following the system reconfiguration, must be used.

In addition to the performance gain in case of a successful retry, the characteristic of a fault can be monitored through retries. For instance, a retry which succeeds after the retry duration  $r'$  implies that the active duration of the fault is also less than or equal to  $r'$ .<sup>2</sup> Even when the retry fails, it indicates that the active duration of the fault is greater than  $r'$ . On the other hand, the detection of fault gives information regarding the duration of fault occurrence and the benign duration of an intermittent fault. Thus, it becomes possible to observe the nature of fault through both retry and detection mechanisms. Note, however, that the information obtained from retry is censored, since for example, in case of an unsuccessful retry, the sampling via retry is stopped while the associated fault is still active. With the censored information, the problem of estimating the nature of fault is the same as the design of experiments in the sequential analysis where the experiments are described by the retry policy, and the sampling is analogous to detection and retry.

The paper first presents a brief description of fault models in Section 2. It should be obvious that  $r'$  will depend on fault behavior, and in Section 3, we begin with how to derive it, given the fault characteristics. When quantitative descriptions of fault behavior are hard to come by in the real environments, the combination of retry and detection enables us to observe the fault characteristics, while determining the optimal retry policy. We counter this in Section 4 by showing how to use statistical estimation theory to create a system that learns via retry the fault characteristics as it goes, and therefore becomes increasingly more "optimal" in the sense of minimizing the task completion time. Due to its repetitive reappearances, retry of an intermittent fault is a renewal process. In Section 5, we apply the Bayes sequential decision to fault characterization and

---

<sup>2</sup>This is not really true due to the fault latency. The fault latency [1] which is defined as the interval between the moments of fault occurrence and error generation has no effect on retry. Thus, we simply ignore the fault latency in the consideration of retry.

retry decision. The backward induction for testing hypotheses is also presented as a solution to the sequential decision problem. The paper concludes with Section 6.

In what follows, we will use continuous retry durations instead of the number of retry attempts.<sup>3</sup>

## 2. BASIC MODEL OF FAULTS

Let a *unit* be the smallest hardware system component that the detection mechanisms can distinguish when a fault is detected, e.g., a processor module can be viewed as an assembly of such units. Hence, the term "module" will mean a system component larger than a unit. Typically, a module is formed by a set of units. For each unit, the fault's behavior can be modeled by a three-state stochastic process as in Figure 1. Denote these three states, namely non-faulty, fault-active, and fault-benign by  $NF$ ,  $F$  and  $FB$  (see [4] for their detailed descriptions). At  $NF$ , no fault exists in the unit. Transition from  $NF$  to  $F$  indicates the occurrence of a fault. If the fault is intermittent and becomes benign following an active duration, the state of the unit changes to  $FB$ . The unit may move back to  $F$  when this intermittent fault recurs -- this is referred to as the *reappearance* of the intermittent fault. If the fault is transient and disappears, the unit will transfer from  $F$  back to  $NF$ . The model similar to this has been widely used in the reliability analyses and the modeling of faults [4,9,10].

Let  $T_f$ ,  $T_i^a$ ,  $T_i^b$  and  $T_i^b$  denote the duration between two successive fault occurrences, the active duration of a transient fault, and the active and benign durations of an intermittent fault, respectively. These durations are random variables with distributions  $F_f$ ,  $F_i^a$ ,  $F_i^a$  and  $F_i^b$ , and density functions  $f_f$ ,  $f_i^a$ ,  $f_i^a$  and  $f_i^b$ , respectively. For simplicity, we

<sup>3</sup>Conversion between a retry duration and its corresponding number of retry attempts is not difficult as discussed in Conclusion.

assume that these durations are mutually independent and that the causes of triggering different types of fault are not correlated. The latter assumption implies that the occurrence of any type of fault can be modeled as a Bernoulli process with probabilities  $p_t$ ,  $p_i$ , and  $p_p$  for transient, intermittent and permanent faults, respectively. Thus, the characteristics of a fault can be represented by a 7-tuple  $C_f \in \{(p_t, p_i, p_p, F_f, F_t^a, F_t^i, F_t^p) \mid p_t + p_i + p_p = 1\}$ .

Usually, the mean time between the occurrences of fault,  $E[T_f]$ , is much larger than any other durations. Thus, it is reasonably accurate to assume that there is at most one fault in a given unit at any moment. In addition, it is assumed that the reappearance of an intermittent fault is never mistaken for the occurrence of a new fault within a unit.<sup>4</sup> Following a successful retry, the detection mechanisms should be able to recognize the type of fault in a unit by continuously monitoring normal operation. When the detection mechanisms find the same unit failing again within a short period, the unit is declared to have an intermittent fault. If the fault has disappeared for a long period, it is regarded as a transient fault.

### 3. OPTIMAL RETRY POLICY FOR GIVEN $C_f$

#### 3.1. General Problem Statement

Once a fault is detected, it is necessary to take a proper sequence of actions such as fault isolation, system reconfiguration, and recovery. For convenience, define the *recovery overhead* as the total time required to resume the normal system operation in case of the detection of a fault; this is a *system-oriented view*. On the other hand, the occurrence of fault may delay the completion of the executing task; this is a *task-*

---

<sup>4</sup>This is the very reason why the term "unit" is introduced here.

*oriented view.* These two views are equivalent when the fault is transient or permanent. For an intermittent fault, since it appears and disappears repetitively, the accumulated overhead of retry could become unacceptably large (eventually infinite). In view of this fact, an intermittent fault has the same undesirable effects on computer performance as a permanent fault when retry is used as the sole means of recovery. Consequently, as far as the minimization of the expected recovery overhead (i.e. the system-oriented view) is concerned, an intermittent fault can be regarded as a permanent fault and hence retry should *not* be used when detection mechanisms find *again* the same fault that was detected but became inactive during the last retry. Once intermittent faults are treated just like permanent faults, the optimal retry policy of minimizing the total recovery overhead becomes equivalent to a special retry policy (for transient faults) which minimizes the task completion time. More on this will be discussed near the end of this subsection.

A most attractive gain from retry is the rescuing of the executing task, i.e., the task-oriented view of retry. Suppose there is enough redundancy so that the system may be reconfigured and the affected task may be migrated to other fault-free modules when a module becomes faulty (due to one or more faulty units within the module). It is obvious that no task should be started on any faulty or potentially faulty module (having one or more units with benign intermittent fault(s)). Consider a practical case in which a module (i) becomes faulty once and gets back to normal during execution of a task, and (ii) never becomes faulty again before the task is completed. In such a case, it is possible to avoid the overhead of migrating and restarting the task by means of a successful retry, leading to a fast completion of the task. Even if the fault that occurred was intermittent, retry is the best recovery method when its active duration is short and benign duration is long, insofar as the completion time of the executing task is

concerned.

Considering the task-oriented view of retry, we will derive an optimal retry policy for a given fault characteristic,  $C_f$ , which minimizes the expected task completion time when a fault is detected during the task execution. Such an optimal policy would be of significant value to real-time applications where small task completion times are of paramount importance.

Let  $x_0$  denote the computation time initially needed to complete the task under a fault-free condition. When a fault is detected, the amount of computation remaining to complete the task, i.e., *residual computation*, is denoted by  $x$ , where  $0 < x < x_0$ . For the  $n$ -th detection of the same fault when the residual computation  $x$  and the characteristic  $C_f$  are both given, the optimal retry policy should specify a *maximum allowable retry duration*,  $r_n^*(x, C_f)$ . When  $n=1$ , the detected fault may be transient, intermittent, or permanent, since the fault type is unknown; but it is intermittent if  $n \geq 2$ . Since the optimal retry policy is to minimize the time necessary to complete the residual computation  $x$  regardless of what has happened in the past, we have the following lemma.

**Lemma 1:**  $r_i^*(x, C_f) = r_j^*(x, C_f)$  for all  $i, j \geq 2$ .

Thus, we have to consider two maximum retry durations for two different cases: the case when a new fault is detected, and the case when an old intermittent fault is detected again. Let  $R = \{(r_1(x, C_f), r_2(x, C_f)) \mid 0 < x < x_0\}$  be a retry policy where the maximum retry durations is  $r_1(x, C_f)$  or  $r_2(x, C_f)$  for the detection of a new fault or an old intermittent fault with the residual computation  $x$  and fault characteristic  $C_f$ . For notational simplicity, we shall use  $r_i$ , whenever convenient, in the sequel, to represent  $r_i(x, C_f)$ ,  $i=1,2$ . Also, denote the expected times needed to complete the residual compu-

tation  $x$  by  $V_1(x, C_f, R)$ ,  $V_2(x, C_f, R)$ ,  $V_3(x, C_f, R)$ , and  $V_4(x, C_f, R)$  when the system is in the following situations: execution starts/resumes on a non-faulty module, a new fault is detected, an old intermittent fault is detected again, and execution continues following a successful retry for an intermittent fault, respectively. Based on transitions among these situations, one can derive the following recursive equations:

$$V_1(x, C_f, R) = \{1 - F_f(x)\}x + \int_0^x \{t + V_2(x-t, C_f, R)\} dF_f(t) \quad (1)$$

$$V_2(x, C_f, R) = p_i \int_0^{r_1} \{t + V_1(x, C_f, R)\} dF_i^a(t) + p_i \int_0^{r_1} \{t + V_4(x, C_f, R)\} dF_i^a(t) \\ + \{1 - p_i F_i^a(r_1) - p_i F_i^a(r_1)\} \{V_1(a(x), C_f, R) + r_1 + t_s\} \quad (2)$$

$$V_3(x, C_f, R) = \{1 - F_i^a(r_2)\} \{V_1(a(x), C_f, R) + r_2 + t_s\} + \int_0^{r_2} \{t + V_4(x, C_f, R)\} dF_i^a(t) \quad (3)$$

$$V_4(x, C_f, R) = \{1 - F_i^b(x)\}x + \int_0^x \{t + V_3(x-t, C_f, R)\} dF_i^b(t) \quad (4)$$

where  $a(x)$  is the residual computation needed when the system applies recovery methods other than retry, e.g.,  $a(x) = x_0$  if restart is used following an unsuccessful retry, and  $t_s$  is the set-up time necessary for system reconfiguration and re-initialization. The optimal retry policy,  $R^* = \{(r_1^*(x, C_f), r_2^*(x, C_f)) \mid 0 < x < x_0\}$ , should minimize both  $V_2(x, C_f, R)$  and  $V_3(x, C_f, R)$  for all  $x$ , since retry is directly applicable only to the second and third situations. (As such, they are explicitly dependent on  $r_1$  and  $r_2$ .) Obviously, this policy also minimizes  $V_1(x, C_f, R)$  and  $V_4(x, C_f, R)$ .

Since the mean time between failures is usually much longer than the other durations,  $V_1(x, C_f, R)$  can be accurately approximated by  $x$ . In general, there are no closed form solutions for  $r_1^*(x, C_f)$  and  $r_2^*(x, C_f)$ . However, these optimal retry durations can be calculated numerically without difficulty as explained below. With the initial condition  $V_4(0, C_f, R) = 0$ ,  $V_3(x, C_f, R)$  and  $V_4(x, C_f, R)$  can be calculated iteratively using Eqs. (3) and

(4) for any given  $R$ . Thus, one can numerically determine  $r_2^*(x, C_f)$  so as to minimize  $V_3(x, C_f, R)$ ;  $V_4(x, C_f, R)$  is also determined. Once  $V_4(x, C_f, R)$  is known, one can easily compute  $V_2(x, C_f, R)$  and therefore  $r_1^*(x, C_f)$ .

When the recovery overhead in place of the task completion time is to be minimized,  $r_2^*(x, C_f) = 0$  for all  $x \in (0, x_0)$ . In this case, the recovery overhead can be expressed as  $V_2(x, C_f, R) - V_1(x, C_f, R)$ , which is the time spent to restore the system to its state immediately before the fault is detected. The optimal retry duration  $r_1^*(x, C_f)$  can be determined through Eqs. (1)-(4) just as we can compute that for minimizing the task completion time. Consequently, we will in the sequel deal with the task completion time only.

### 3.2. Fault Active Durations with Monotone Hazard Rate Functions

Since  $T_f^a$  is a continuous random variable, one can assume that  $f_1^a(t)$  is continuous in  $[0, \infty)$ . The hazard rate function of the active duration of an intermittent fault is defined by  $h_1^a(t) \equiv \frac{f_1^a(t)}{1 - F_1^a(t)}$ . When the hazard rate function of the active duration of an intermittent fault is monotonically increasing, constant, or monotonically decreasing, the optimal retry duration  $r_2^*$  exhibits interesting properties. These properties play a significant role in determining the optimal retry policy, since the time durations associated with faults are usually modeled to have monotone hazard rate functions. Typical distributions with monotonically increasing hazard rate functions include the gamma and the Weibull distributions with the shape parameters greater than 1. When their shape parameters are less than 1, they have monotonically decreasing hazard rate functions. The exponential distribution has a constant hazard rate. Consider first the non-decreasing hazard rate function which leads to the following theorem.

**Theorem 1:** When  $h_i^a(t)$  is monotonically non-decreasing in  $t$ ,  $r_2^* = 0$  or  $r_2^* = \infty$ .

*Proof:* Differentiating Eq. (3) with respect to  $r_2$ , we obtain

$$\begin{aligned} \frac{\partial V_3(x, C_f, R)}{\partial r_2} &= f_i^a(r_2) \left[ V_4(x, C_f, R) + \frac{1}{h_i^a(r_2)} - \{V_1(a(x), C_f, R) + t_s\} \right] \\ &\stackrel{\circ}{=} f_i^a(r_2) \left[ V_4(x, C_f, R) + \frac{1}{h_i^a(r_2)} - \{a(x) + t_s\} \right] \end{aligned} \quad (5)$$

Since  $V_4(x, C_f, R)$  is independent of the past and current retry durations  $r_2(y, C_f)$  where  $y \geq x$ ,<sup>5</sup>  $V_4(x, C_f, R) + \frac{1}{h_i^a(r_2)} - \{a(x) + t_s\}$  is non-increasing in  $r_2(x, C_f)$ . If there is an  $r$  such that  $V_4(x, C_f, R) + \frac{1}{h_i^a(r)} - \{a(x) + t_s\} < 0$ , then the first derivative of  $V_3(x, C_f, R)$  with respect to  $r_2$  is negative for all  $r_2 > r$ . Thus,  $r_2^* = \infty$ . If such an  $r$  does not exist, the derivative is always non-negative, implying that  $V_3(x, C_f, R)$  is monotonically non-decreasing. This results in  $r_2^* = 0$ . Q.E.D.

Following the definition of  $r_k^i(x, C_f)$ ,  $r_2(x, C_f) = 0$  implies that no retry be attempted for reappearing intermittent faults, whereas  $r_2^i(x, C_f) = \infty$  means that the retry should be applied until the intermittent fault becomes benign.

**Corollary 1:** When  $h_i^a(t)$  is monotonically non-decreasing in  $t$  and if there exists an  $x_2^*$  such that  $a(x_2^*) + t_s - x_2^* - (R_i^b(x_2^*) - 1)E[T_i^a] = 0$  where  $R_i^b(x)$  is the renewal function [11] corresponding to the distribution  $F_i^b(t)$ , then  $r_2^i(x, C_f) = \infty$  if  $x \leq x_2^*$  and  $r_2^i(x, C_f) = 0$  otherwise.

*Proof:* From Theorem 1,  $r_2^i(x, C_f)$  is either  $\infty$  or 0. When  $r_2^i(x, C_f) = \infty$ , there exists an  $r$

<sup>5</sup> Note that the probability of having a zero benign duration of an intermittent fault should be zero, i.e.  $Pr.b(T_i^b = 0) = 0$ . Otherwise, no useful computation can be done.

such that Eq. (5) becomes negative. Since  $V_4(x, C_f, R^q)$  is monotonically non-decreasing functions of  $x$ , there also exists an  $r$  such that Eq. (5) becomes negative when the residual computation is less than  $x$ . Thus,  $r_2^q(y, C_f) = \infty$  for all  $y \leq x$ . Since both the active and benign durations are mutually independent, we have

$$V_4(x, C_f, R^q) = x + \{E[N(x)] - 1\}E\{T_1^q\}$$

where  $N(x)$  is the number of reappearances of the intermittent fault during the residual computation  $x$ , namely  $N(x) = \inf \{n; \sum_{k=1}^n T_{i,k}^b \geq x\}$  where  $T_{i,k}^b$  is the benign duration following the  $k$ -th occurrence of the intermittent fault. The expected value of  $N(x)$ ,  $E[N(x)]$ , is equivalent to the renewal function  $R_1^b(x)$  corresponding to the distribution  $F_1^b(t)$ . Also,  $r_2^q(x, C_f) = \infty$  if and only if  $V_3(x, C_f, R)|_{r_2=\infty} \leq V_3(x, C_f, R)|_{r_2=0}$ , i.e.,

$$\int_0^\infty (t + V_4(x, C_f, R^q)) dF_1^q(t) = E\{T_1^q\} + V_4(x, C_f, R^q) \leq a(x) + t_s$$

From the equality in the right-hand side of the above equation, we obtain  $x_2^q$  and thus the Corollary is proved. Q.E.D.

Theorem 1 can also be viewed as below using the concept of *stochastic ordering* between two random variables. A random variable  $X$  is said to be *stochastically larger than* the other random variable  $Y$  if  $Prob(X > t) \geq Prob(Y > t)$  for all  $t$  [12]. Let  $T_1^q(r)$  be the remaining life of the intermittent fault after retry has been applied for the duration  $r$ . When the hazard rate function is non-decreasing,  $T_1^q(r)$  is stochastically larger than  $T_1^q(s)$  provided  $r \leq s$ . Thus, for all  $s \geq r$ , if it is worth continuing retry beyond the retry duration  $r$  (in the sense of minimizing the task completion time), then we should continue the retry even after the retry duration  $s$ . Consequently, the retry continues until the intermittent fault disappears.

Note that when the hazard rate function is non-decreasing,  $x_2^*$  is determined by the mean active duration and is independent of the shape of the distribution.  $x_2^*$  could become negative when  $E[T_i^a]$  is large, that is, intermittent faults have a long active duration. In such a case, Corollary 1 implies that no retry be applied for intermittent faults. On the other hand, if the set-up overhead  $t_s$  is large,  $x_2^*$  could be even larger than  $x_0$ , implying that retry be used as a sole means of recovering from an intermittent fault.

When the hazard rate  $h_i^a(t)$  is decreasing, the nice properties stated in both Theorem 1 and Corollary 1 do not exist. However, there exists at most one root of Eq. (5) that minimizes  $V_3$ . In such a case, since there is no closed form expression of  $V_4(x, C_f, R')$ , we have to resort to (less elegant) numerical techniques for determining both  $r_2^*(x, C_f)$  and  $r_1^*(x, C_f)$  as was previously mentioned.

Several numerical examples, in which restart is used as a sole means of backup recovery, i.e.  $a(x) = x_0$ , are shown in Figures 2 to 4. In these figures, the durations are normalized with respect to  $x_0$ , and the active duration of the intermittent fault is assumed to have the gamma or Weibull distribution. Figure 2 presents  $x_2^*$ 's when the shape parameters  $\alpha$ 's of the gamma and Weibull distributions, respectively, are greater than or equal to 1. Figures 3 and 4 show the optimal retry duration  $r_2^*(x, C_f)$ ; the solid lines for  $\alpha < 1$  and the dashed lines for  $\alpha = 1$ . Note that for the gamma distribution  $h_i^a(t)$  approaches  $\frac{1}{\beta}$  as  $t \rightarrow \infty$  where  $\beta$  is the scale parameter. Thus, it is possible for the derivative of  $V_3$  to be negative, (i.e. Eq. (5) becomes negative), implying  $r_2^*(x, C_f) = \infty$ . For the Weibull distribution with  $\alpha < 1$ ,  $r_2^*$  never becomes  $\infty$  since  $h_i^a(\infty) = 0$ .

Consider the case where  $T_t$ ,  $T_i^a$  and  $T_i^b$  are all exponentially distributed with the parameters  $\tau$ ,  $\mu$ ,  $\nu$  for the transient fault disappearance rate, the intermittent fault disappearance and reappearance rates, respectively. Since  $f_i^b(t) = \nu e^{-\nu t}$ , the renewal

function  $R_1^*(x)$  becomes  $1+\nu x$ . From Corollary 1, we have  $r_2^*(x, C_f) = \infty$  if  $x \leq x_2^*$  and  $r_2^*(x, C_f) = 0$  if  $x > x_2^*$ , where  $x_2^* = \frac{\mu}{\mu+\nu} (x_0 + t_r - \frac{1}{\mu})$ . Since  $V_4(x, C_f, R)$  implicitly depends on  $r_2$  via  $V_3$ , we can express  $V_4(x, C_f, R')$  as:

$$V_4(x, C_f, R') = \begin{cases} (1 + \frac{\nu}{\mu})x & \text{if } x \leq x_2^* \\ x_0 + t_r + \frac{1}{\nu} - e^{-\nu(x-t_r)}(\frac{1}{\mu} + \frac{1}{\nu}) & \text{if } x > x_2^* \end{cases} \quad (8)$$

The derivative of  $V_2(x, C_f, R)$  with respect to  $r_1$  becomes:

$$\frac{\partial V_2(x, C_f, R)}{\partial r_1} = p_p + p_i e^{-r_1} \{1 - (x_0 + t_r - x)r\} + p_r e^{-r_1} \{1 - (x_0 + t_r - V_4(x, C_f, R))\mu\} \quad (7)$$

With  $r_2^*(x, C_f)$  as determined in Corollary 1 and  $V_4(x, C_f, R')$  as in Eq. (8), Eq. (7) can have at most two roots. The optimal retry duration  $r_1^*(x, C_f)$  can be obtained by examining  $V_2(x, C_f, R)$  at the boundaries,  $r_1=0$  and  $r_1=\infty$ , and the roots of Eq. (7). Note that  $r_1^*$  cannot be infinite as long as  $p_p > 0$ . Unlike  $r_2^*$ ,  $r_1^*$  does not have to be zero when  $x > x_2^*$ . Several cases of  $V_2(x, C_f, R)$  as a function of  $r_1$  are shown in Figure 5 where all parameters are normalized with respect to  $x_0$ . The case 2 in Figure 5 shows an example for which two positive roots of Eq. (7) exist. Figure 6 presents some numerical results on  $r_1^*(x, C_f)$  as a function of  $x$ . Note that  $x_2^*$  depends upon the ratio of  $\nu$  to  $\mu$ , whereas  $r_1^*$  varies as  $p_i$ ,  $p_r$  and  $p_p$  change.

#### 4. OPTIMAL RETRY POLICY AND PARAMETER ESTIMATION

In Section 3, we have derived an optimal retry policy for a given fault characteristic  $C_f$ . It is, however, very difficult in practice to know *a priori* the fault characteristic. Even if the fault characteristic is measured during device manufacture, it may well vary as the execution environment and the executing tasks change. Another factor that

makes the fault characteristics time-variant is the aging of components, e.g., the bathtub curve of the failure rate as a function of time [1]. Thus, it is important to determine an optimal retry policy for uncertain fault characteristics. Note that retry not only provides an efficient recovery of task execution, but also monitors the behavior of a fault present in a hardware unit. Naturally, it is desirable to integrate the estimation of the fault characteristics and the control of the maximum retry durations into a single decision problem. In such a case, the computer system has to adjust its retry policy using the information on the fault behavior collected during its past retries.

The detection mechanisms<sup>6</sup> can be useful in estimating the duration between two successive fault occurrences or the benign duration of an intermittent fault. Note that this information is crucial in specifying the behavior of fault occurrence or reappearance. Consequently, the fault characteristics would become well-defined if a good estimator were used. Moreover, retry may lead to an indication of the active duration of a transient or intermittent fault, which is, on the other hand, affected by the retry policy applied. The information collected is incomplete in case of an unsuccessful retry, since the retry is stopped while the associated fault is active. In what follows, we consider the estimation of the characteristic of an active fault and the simultaneous determination of an optimal retry policy which minimizes the task completion time.

Note that the probabilities of having a permanent, transient, or intermittent fault are crucial to the determination of  $r_1'(x, C_f)$  but unrelated to that of  $r_2'(x, C_f)$ . It implies that correlations among successive retry durations during the execution of a task do not depend on these probabilities. Thus, to minimize the task completion time, it is assumed that these probabilities are determined *a priori* from the previous observations

---

<sup>6</sup>As was pointed out, we mean here the signal-level detection of faults [4].

of fault occurrence. These probabilities can be estimated accurately if a sufficiently large amount of data of fault occurrence were collected. If a sufficient number of samples has not yet been obtained, the measured results as in [2,3] have to be used instead.

Recall that in the determination of  $r_2^i$ , the transient fault is excluded. Also, if  $p_0$ ,  $p_D$ ,  $p_p$  are determined *a priori*, the effects of retry on the task completion time is a linear combination of the effects of transient and intermittent faults when a new fault is detected. This implies that the same technique can essentially be used to deal with the unknown parameters for both transient and intermittent faults. Consequently, we confine ourselves to the case where the density function of the active durations of transient faults is known and the active durations of intermittent faults have the density function form  $f_i^*(t|\theta)$  with the parameter  $\theta$  unknown. (Note that  $\theta$  could be a vector if there are two or more parameters, e.g., the shape parameter  $\alpha$  and the scale parameter  $\beta$  for the Weibull and gamma distributions.)

The samples obtained from retry can be represented by a 2-tuple  $(I, t)$  where  $I$  is a single-bit flag and  $t$  indicates a duration.  $I=0$  represents a successful retry, and hence  $t$  indicates the active duration of the fault. On the other hand, when a retry fails,  $I=1$  and  $t$  is the retry duration. Let  $(I_i, t_i)$   $i=1,2,\dots,n$  denote the past samples related to the active duration of an intermittent fault. These resulting samples are *type I progressively censored*, following Cohen's definition in [13] with continuous censoring times. There are several different types of estimators conceivable for estimating the parameter  $\theta$  on the basis of these progressively censored samples. For the Weibull and gamma distributions, the maximum likelihood estimators have been widely studied as in [13-17] when the samples are progressively censored. For simplicity (but not because of difficulty), we shall employ the maximum likelihood estimator  $\hat{\theta}$  of  $\theta$  in the sequel.

When the fault is still active even after the current retry has been applied for the duration  $r$ , we shall have collected an additional sample  $(1, r)$  via the current retry. Let  $\hat{\theta}(r)$  be the maximum likelihood estimator of  $\theta$  which is based on the samples including up to the current sample  $(1, r)$ . Following the current retry duration  $r$ , the maximum likelihood function becomes

$$L(\theta) = \left\{ \prod_{i=1}^n \eta(I_i, t_i, \theta) \right\} \eta(1, r, \theta) \quad (8)$$

where  $\eta(I, t, \theta)$  is defined as:

$$\eta(I, t, \theta) = \begin{cases} f_i^q(t|\theta) & \text{if } I=0 \\ 1-F_i^q(t|\theta) & \text{if } I=1 \end{cases}$$

The maximum likelihood estimator  $\hat{\theta}(r)$  should maximize  $L(\theta)$  or  $\log L(\theta)$ .

Let the optimal retry durations based on the estimated  $\hat{\theta}(r)$  be denoted by  $r_k^*(x, \hat{\theta}(r))$   $k=1,2$  for a newly detected fault and an old intermittent fault, respectively. Use the notation  $C_f(\hat{\theta}(r))$  to indicate that the active durations of intermittent faults have the density function  $f_i^q(t|\hat{\theta}(r))$ , and let  $R(r)$  denote the policy that the maximum allowable retry duration for the current retry is  $r$ . Then, the direct solution of the optimal retry duration is to find the minimum of  $V_k(x, C_f(\hat{\theta}(r)), R(r))$   $k=1,2,3,4$ . Notice that the retry duration  $r$  not only appears in the integral equations (2) and (3), but also affects the fault characteristic  $C_f$ .

Under certain conditions, it can be proven that  $r_k^*(x, \hat{\theta}(r))$  is a non-increasing function of  $r$ . We will first derive the results under such conditions, the application of which to a more general case is then discussed later in this section. For the former case, the optimal retry duration  $r_k^*$  for the current retry can be readily obtained by the following

theorem.

**Theorem 2:** When (a) the active duration of an intermittent fault has the density function  $f_i^a(t|\theta)$ , and (b) for  $t_j \geq t_k$  the ratio  $\frac{f_i^a(t|\hat{\theta}(t_j))}{f_i^a(t|\hat{\theta}(t_k))}$  -- a non-decreasing likelihood ratio

[18] -- is non-decreasing in  $t$ , then the optimal retry duration is determined by:

$$r_k^* = \inf \{ r; \hat{r}_k^*(x, \hat{\theta}(r)) \leq r \} \quad (9)$$

To prove this theorem, we need the following three lemmas.

**Lemma 2.** Under the same conditions as in Theorem 2, let  $T_j$  and  $T_k$  be random variables with the density functions  $f_i^a(t|\hat{\theta}(t_j))$  and  $f_i^a(t|\hat{\theta}(t_k))$ , respectively, and  $\Psi(t)$  be a non-decreasing function of  $t$ , then  $E[\Psi(T_j)] \geq E[\Psi(T_k)]$  provided  $t_j \geq t_k$ .

Proof of this lemma follows immediately from Lemma 2 of Chapter 3 in [18].

Let  $h_i^a(t|\hat{\theta}(t_j))$  be the hazard rate function when the density function of  $T_j^a$  is  $f_i^a(t|\hat{\theta}(t_j))$ . The following Lemma gives the ordering of  $h_i^a(t|\hat{\theta}(t_j))$  with respect to  $t_j$ .

**Lemma 3.** Under the same conditions as in Theorem 2,  $h_i^a(t|\hat{\theta}(t_j))$  is a non-decreasing function of  $t_j$  for every fixed  $t$ .

*Proof:* For  $t_j \geq t_k$ , we have  $\frac{f_i^a(t|\hat{\theta}(t_j))}{f_i^a(t|\hat{\theta}(t_k))} \leq \frac{f_i^a(s|\hat{\theta}(t_j))}{f_i^a(s|\hat{\theta}(t_k))}$  for all  $s \geq t$ . This inequality implies

that  $\frac{f_i^a(t|\hat{\theta}(t_j))}{f_i^a(t|\hat{\theta}(t_k))} \leq \frac{\int_t^\infty f_i^a(u|\hat{\theta}(t_j)) du}{\int_t^\infty f_i^a(u|\hat{\theta}(t_k)) du} = \frac{1-F_i^a(t|\hat{\theta}(t_j))}{1-F_i^a(t|\hat{\theta}(t_k))}$ . Thus,  $h_i^a(t|\hat{\theta}(t_j)) \geq h_i^a(t|\hat{\theta}(t_k))$  if  $t_j \geq t_k$ .

Q.E.D.

Let  $V_k'(x, \hat{\theta}(t)) = \min_R V_k(x, \hat{\theta}(t), R)$   $k=2,3,4$  where  $\hat{\theta}(t)$  is used in place of  $C(\hat{\theta}(t))$ .

Note in this case that the active duration of the intermittent fault is distributed with the parameter  $\hat{\theta}(t)$  and that all the other distributions are known.

**Lemma 4.** Under the same conditions defined as in Theorem 2, if  $t_1 > t_2$ , then

$$(i) V_k'(x, \hat{\theta}(t_1)) \geq V_k'(x, \hat{\theta}(t_2)) \quad k=2,3,4,$$

$$(ii) r_k'(x, \hat{\theta}(t_1)) \leq r_k'(x, \hat{\theta}(t_2)) \quad \text{for } k=1,2.$$

*Proof:* The proof for  $k=3,4$  is done by mathematical induction. Let  $V_{k,n}(x, \hat{\theta}(t_j), r_2(n, j))$   $k=3,4$  be the expected times needed to complete the residual computation  $x$  when there are at most  $n$  retries to be attempted following the current one, and let  $r_2(n, j)$  be the maximum retry duration allowed. Also, let the optimal retry duration to achieve the minimum  $V_{k,n}(x, \hat{\theta}(t_j))$  be  $r_2'(n, j)$ . For  $n=0$ ,  $V_{4,0}(x, \hat{\theta}(t_j))=x$  and

$$V_{3,0}(x, \hat{\theta}(t_1)) - V_{3,0}(x, \hat{\theta}(t_2), r_2'(0,1)) = \int_0^\infty \Psi(t, x, r_2'(0,1)) \{f_1'(t|\hat{\theta}(t_1)) - f_1'(t|\hat{\theta}(t_2))\} dt$$

where  $\Psi(t, x, y) = t+x$  when  $t \leq y$  and  $y+x_0+t$ , when  $t > y$ . Since  $\Psi(t, x, r_1^0)$  is non-decreasing in  $t$ , the right-hand side of the above equation is non-negative as a result of Lemma 2. Also, since  $V_{3,0}(x, \hat{\theta}(t_2))$  is the minimum when the active duration of the intermittent fault has the density function  $f_1'(t|\hat{\theta}(t_2))$ , then

$$V_{3,0}(x, \hat{\theta}(t_1)) \geq V_{3,0}(x, \hat{\theta}(t_1), r_2'(0,1)) \geq V_{3,0}(x, \hat{\theta}(t_2)).$$

Suppose that  $V_{3,n}(x, \hat{\theta}(t_1)) \geq V_{3,n}(x, \hat{\theta}(t_2))$  and  $V_{4,n}(x, \hat{\theta}(t_1)) \geq V_{4,n}(x, \hat{\theta}(t_2))$  for all  $x$  provided  $t_1 \geq t_2$ . It is obvious to see from Eq. (4) that  $V_{4,n+1}(x, \hat{\theta}(t_1)) \geq V_{4,n+1}(x, \hat{\theta}(t_2))$  for all  $x$ . Thus,

$$V_{3,n+1}(x, \hat{\theta}(t_1)) - V_{3,n+1}(x, \hat{\theta}(t_2), r_2'(n+1,1)) \geq \int_0^\infty \Psi(t, V_{4,n+1}(x, \hat{\theta}(t_1)), r_2'(n+1,1)) \{f_1'(t|\hat{\theta}(t_1)) - f_1'(t|\hat{\theta}(t_2))\} dt$$

$\Psi(t, V_{i,n+1}^*(x, \hat{\theta}(t_1)), r_2^*(n+1, 1))$  is non-decreasing in  $t \leq r_2^*(n+1, 1)$ . Also, since  $r_2^*(n+1, 1)$  is the optimal retry duration,  $V_{i,n+1}^*(x, \hat{\theta}(t_1)) \leq x_0 + t_r$ . Hence,  $\Psi(t, V_{i,n+1}^*(x, \hat{\theta}(t_1)), r_2^*(n+1, 1))$  is always non-decreasing in  $t$ . The right-hand side of the above equation becomes non-negative, resulting in  $V_{3,n+1}^*(x, \hat{\theta}(t_1)) \geq V_{3,n+1}^*(x, \hat{\theta}(t_2), r_2^*(n+1, 1)) \geq V_{3,n+1}^*(x, \hat{\theta}(t_2))$ . By mathematical induction, we have  $V_k^*(x, \hat{\theta}(t_1)) \geq V_k^*(x, \hat{\theta}(t_2))$  for  $k=3, 4$ .

To prove  $r_2^*(x, \hat{\theta}(t_1)) \leq r_2^*(x, \hat{\theta}(t_2))$ , the following cases are examined. When  $r_2^*(x, \hat{\theta}(t_1)) = 0$ , the relation is always true. When  $r_2^*(x, \hat{\theta}(t_1)) > 0$ , using Lemma 3 and the first part of this proof, the derivative of  $V_3(x, \hat{\theta}(t_j), R)$  with respect to the retry duration  $r$  has the following ordering relationship for all  $r$  and  $t_1 \geq t_2$ .

$$V_4^*(x, \hat{\theta}(t_1)) + \frac{1}{h_1^2(r|\hat{\theta}(t_1))} - \{a(x) + t_r\} \geq V_4^*(x, \hat{\theta}(t_2)) + \frac{1}{h_1^2(r|\hat{\theta}(t_2))} - \{a(x) + t_r\}$$

where all retries after the current one are assumed to employ the optimal policy. Thus, for  $t_1 \geq t_2$ ,  $r_2(x, \hat{\theta}(t_2)) = \infty$  when  $r_2^*(x, \hat{\theta}(t_1)) = \infty$ , and  $r_2(x, \hat{\theta}(t_2)) \geq r_2^*(x, \hat{\theta}(t_1))$  when  $r_2^*(x, \hat{\theta}(t_1))$  is finite.

For the case of  $k=2$ , it is easy to see that  $V_2(x, \hat{\theta}(t_j), R)$  is a linear combination of the effects of both transient and intermittent faults. Thus,  $V_2(x, \hat{\theta}(t_j), R^*) \geq V_2(x, \hat{\theta}(t_k), R^*)$ . Also, the handling of  $V_2$  with respect to  $r_1$  has the same ordering relationship as that of  $V_3$  with respect to  $r_2$ . Thus,  $V_2(x, \hat{\theta}(t_j), R^*) \geq V_2(x, \hat{\theta}(t_k), R^*)$ , and  $r_1^*(\hat{\theta}(t_j)) \leq r_1^*(\hat{\theta}(t_k))$  when  $t_j \geq t_k$ . Q.E.D.

Lemma 4 shows that  $r_k^*(x, \hat{\theta}(t_j))$  is non-increasing in  $t_j$  for  $k=1, 2$ . Thus, there exists an  $r$  such that  $r \geq r_k^*(x, \hat{\theta}(r))$ . The proof of Theorem 2 is given as follows:

*Proof of Theorem 2:* Suppose that the retry has been applied for the period  $r$  but the fault is still active. When  $r_k^*(x, \hat{\theta}(r)) > r$ , the retry should be continued since it decreases

the expected task completion time. Thus,  $r_k'(x) > \sup \{r; r_k'(x, \hat{\theta}(r)) > r\}$ . Suppose there is an  $r_j \in \{r; r_k'(x, \hat{\theta}(r)) \leq r\}$ . Then  $V_k(x, \hat{\theta}(r_j), r_j) \geq V_k^*(x, \hat{\theta}(r_j)) \geq V_k^*(x, \hat{\theta}(r_k))$  where  $r_k^*$  is defined in Eq. (9) and  $k' = k+1$ . Thus, the theorem follows. Q.E.D.

For the same example in Section 3.2, suppose the active duration of an intermittent fault is exponentially distributed with an unknown disappearance rate  $\mu$ . Using a method similar to the Cohen's derivation in [13], the maximum likelihood estimator  $\hat{\mu}(r)$  for an exponential distribution -- which maximizes  $\log L(\mu)$  -- is obtained as:

$$\hat{\mu}(r) = (n - \sum_{i=1}^n I_i) \frac{1}{\sum_{i=1}^n t_i + r} \quad (10)$$

Theorem 2 gives the optimal stopping time for the current retry. Note that the true value of  $\mu$  is unknown and its maximum likelihood estimator is to determine the optimal retry duration. In the case of retry for a reappearing intermittent fault, the optimal retry duration for a given  $\mu$  is either 0 or  $\infty$  as shown in Corollary 1. Using Theorem 2 and Eq. (10), we get the optimal retry duration as follows:

$$r_2^* = \max \left[ 0, \left\{ (n - \sum_{i=1}^n I_i) \frac{x_0 + t_0 - x}{1 + \nu x} - \sum_{i=1}^n t_i \right\} \right] \quad (11)$$

Note that the gamma distribution has a non-decreasing likelihood ratio for both  $\alpha$  and  $\beta$  [18]. Furthermore, the estimators provided by Cohen [15] show that both the estimated  $\alpha$  and  $\beta$  are increasing in the current retry period  $r$ . Thus, Theorem 2 can be applied directly when the active duration of the intermittent fault has the gamma distribution. When the distribution of the active duration is Weibull, Theorem 2 cannot be applied directly, but still provides a good approximate solution. This is due to the fact that the Weibull distribution has a non-decreasing likelihood ratio with respect to its

scale parameter only. Since (i) the variation of the estimated shape parameter  $\alpha$  with respect to the current retry duration  $r$  is always less than that of the estimated scale parameter  $\beta$ , and (ii) the estimated  $\beta$  is increasing in  $r$  when  $\alpha$  is fixed, a reasonably good approximation can be obtained by assuming that  $\alpha$  is constant during the current retry and  $\beta$  is estimated using both the past and current samples as discussed in the above.

There are some shortcomings when the maximum likelihood estimator is used for the progressively censored samples. Particularly, the estimator is biased when the samples are censored. Also, in the case of the exponential distribution,  $\hat{\mu}$  does not contain sufficient statistics of  $\mu$  when the samples are censored and incomplete, i.e., when there exists at least one sample  $(I_i, t_i)$  with  $I_i=1$ . These shortcomings can be seen easily from a trivial example:  $\hat{\mu}$  becomes zero when  $I_i=1$  for all  $i=1,2,\dots,n$ . In fact, as shown by van Zwet [19], for most practical cases it is impossible to obtain unbiased estimators when the samples are Type I censored in a semi-infinite interval. Note, however, that there is no restriction about which estimator to be used in the foregoing determination of the optimal retry policy, meaning that estimators other than the maximum likelihood estimator can be used without altering our method described thus far.

## 5. BAYES SEQUENTIAL ANALYSIS AND OPTIMAL RETRY

In the previous section, the unknown parameters of a distribution are estimated first, and the optimal retry policy is then determined using the estimated results. In this section, the same problem is attacked by taking the Bayes approach. Since the reappearances of an intermittent fault during the execution of a task are a renewal process, there could be a sequence of retries for the same intermittent fault. Thus, it is natural

to incorporate the Bayes sequential analysis to both *characterize* the intermittent fault and *determine* the optimal retry policy. Retry is then considered as sampling of the fault behavior, and the retry duration controls both the task completion time and the information to be collected. Since the retry for a newly detected fault only occurs once within the execution of a task (under the assumption that  $E\{T_f\}$  is much larger than the other durations), we focus on the behavior of intermittent faults and thus the determination of  $r_2'$ .

### 5.1. Optimal Retry and Bayes Decision

Let the distribution of  $T_i^q$  be governed by some unknown parameters  $W_i$ . Note that  $W_i$  may be a scalar (e.g. for the exponential distribution) or a vector (e.g. for the gamma or Weibull distribution). The *a priori* information concerning  $W_i$  is expressed in terms of a probability distribution function defined on  $\Omega$ . Let the density function of  $W_i$  be  $\xi_i(w)$ . Denote further the fault characteristics, given  $w_i$  and the prior density function  $\xi_i$ , by  $C_{f,w_i}$  and  $C_{f,\xi_i}$ , respectively.

To apply the Bayes decision theory, the risk with a retry policy  $R$ , given  $\xi_i$  and the residual computation  $x$ , is defined as follows:

$$\rho_k(x, \xi_i, R) \equiv \int_{\Omega} V_k(x, C_{f,w_i}, R) \xi_i(w) dw \quad k=3,4 \quad (12)$$

Thus, the (optimal) *Bayes risk* is given as

$$\rho_k'(x, \xi_i) \equiv \inf_R \rho_k(x, \xi_i, R) \quad k=3,4 \quad (13)$$

Since we are now concerned only with the retry for an intermittent fault,  $R$  consists of  $r_2$  only. The optimal retry duration in case of the detection of an old intermittent fault,  $r_2'(x, C_{f,\xi_i})$ , abbreviated by  $r_2'(x, \xi_i)$ , yields the Bayes risk  $\rho_3'(x, \xi_i)$ . Similarly, the Bayes

risk of the retry for a newly detected fault can be defined with Eqs. (12) and (13). However, the determination of  $r_1'(z, \xi_i)$  is a one stage Bayes decision problem. Once  $\rho_1(z, \xi_i)$  and  $r_2'(z, \xi_i)$  are obtained, the normal form of analysis [22] can be applied directly for the solution of  $r_1'(z, \xi_i)$ .

Following a retry attempt for an intermittent fault, regardless whether it fails or succeeds, an event related to the fault active duration  $T_i^a$  is observed. The event observed during a retry of the duration  $r$  is either "success" or "fail". The "success" event, denoted by  $e'(t)$ , occurs when the detected fault disappears after the retry duration  $t$  which is less than or equal to the maximum allowable retry duration  $r$ . The "fail" event, denoted by  $e'(r)$ , occurs when the detected fault does not disappear by the end of the retry duration  $r$ . Let  $S(r) = \{e'(t); t \leq r\} \cup \{e'(r)\}$ . With the prior density function  $\xi_i(w)$ , the posterior density function following the observation of  $e \in S(r)$ , denoted by  $\xi_i(w|e)$   $i=1,2$ , becomes

$$\xi_i(w|e) = \frac{g(e|w) \xi_i(w)}{\int_{\Omega} g(e|w) \xi_i(w) dw} \quad (14)$$

where  $g(e|w)$  is the generalized conditional density function for the event  $e$  as in [20], i.e.,

$$g(e|w) = \begin{cases} \text{The density function of } T_i^a \text{ at } t & \text{if } e=e'(t) \text{ and } t \leq r \\ \text{Prob}(e'(r)) & e=e'(r) \end{cases} \quad (15)$$

This posterior density function will become the prior density for the next retry. Consequently, the system's behavior is similar to a sequential decision procedure which determines first a retry policy and then observes the resulting sample. The procedure will be repeated with a new prior distribution which is determined on the basis of the new sample observed and the old prior distribution. The decision on retry and the sampling for fault characterization will continue as long as there is an occurrence of fault.

The problem of selecting the optimal retry policy can also be treated as the optimal stopping problem with continuous observations [21]. Suppose an intermittent fault is detected again when the residual computation is  $x$ . Then, retry is applied for a specified stopping time  $r$ . The task will be continued, without applying recovery methods other than retry, if the fault disappears during the retry period  $r$ . Otherwise, it has to be restarted from the beginning.<sup>7</sup> The posterior density function of  $w_i$  becomes  $\xi_i(w|e'(t))$  or  $\xi_i(w|e'(r))$ , depending on the outcome of retry. The cost of an observation is the amount of time used for monitoring the fault until its disappearance, i.e.,  $c(e'(t))=t$ , or until the end of retry, i.e.,  $c(e'(r))=r$ . The costs associated with the termination of retry are defined as the amount of time necessary to complete the residual computation  $x$  as follows:

$$L(x, r, \xi_i | e'(t)) = \rho_i^*(x, \xi_i(w|e'(t)))$$

$$L(x, r, \xi_i | e'(r)) = x_0 + t_0$$

The expected loss for the stopping time  $r_2^*$  is the same as the Bayes risk defined in Eq. (13). According to the theory presented by Irle in [21], there always exists an optimal stopping time,  $r_2^* \in [0, \infty)$ , satisfying Eq. (13).

We will in the next section solve the sequential decision problem using the backward induction [20] for testing hypotheses where the prior and posterior distributions are confined within the open unit interval, i.e., (0,1). Note that the minimax method in [22] cannot be used to solve Eqs. (12) and (13), since the decision space -- which consists of all possible maximum retry durations -- is neither compact nor finite.

---

<sup>7</sup>For simplicity, it is assumed that there is only one alternative to the retry recovery, i.e., restart.

### 5.2. Optimal Retry and Hypotheses Testing

Suppose that there are a primary and some alternative hypotheses concerning the characteristics of an intermittent fault. Consider the sequential testing of these hypotheses and the simultaneous determination of the optimal retry policy; this is not difficult to solve since both the prior and posterior probabilities lie in the same unit interval (0,1). For given hypotheses, the initial prior distribution can be assumed to be equally likely among the hypotheses.

To be more specific, take a demonstrative example<sup>8</sup> in which the active duration of an intermittent fault is assumed to be exponentially distributed with an unknown parameter  $\mu$ . Let there be two hypotheses on  $\mu$ ,  $H_0$  and  $H_1$  for  $\mu = \mu_0$  and  $\mu = \mu_1$ , respectively, and let  $\mu_0 > \mu_1$ . The uncertainty associated with these hypotheses can be represented by the probability  $h$  of having  $\mu = \mu_0$ . We will first determine the optimal retry policy for all  $h \in (0,1)$ . Then, we will consider the problem of testing hypotheses as well as estimating the expected sample size to reach a certain significance level under the optimal retry policy.

Consider the optimal retry duration  $r_2^*(x, h)$  upon detection of an old intermittent fault. In this case, we get the posterior probabilities given the events  $e^*(t)$  and  $e^*(r)$ , denoted by  $h(t)$  and  $\bar{h}(r)$ , respectively as follows:

$$h(t) = \frac{h\mu_0 e^{-\mu_0 t}}{h\mu_0 e^{-\mu_0 t} + (1-h)\mu_1 e^{-\mu_1 t}} \quad \text{where } t \leq r \quad (16)$$

$$\bar{h}(r) = \frac{h e^{-\mu_0 r}}{h e^{-\mu_0 r} + (1-h) e^{-\mu_1 r}} \quad (17)$$

---

<sup>8</sup>As will be pointed out near the end of this section, the results obtained from this example are applicable/extendable to more general cases.

As was discussed in Section 3.2, we can compute  $x_2^*$  for a given  $\mu_i$ , denoted by  $x_2^*(\mu_i)$ ,  $i=0,1$ , such that (i)  $r_2^*(x,1)=\infty$  if  $x \leq x_2^*(\mu_0)$ , or 0 otherwise, and (ii)  $r_2^*(x,0)=\infty$  if  $x \leq x_2^*(\mu_1)$ , or 0 otherwise. Since  $x_2^*(\mu_0) > x_2^*(\mu_1)$ ,  $r_2^*(x,h)=\infty$  if  $x \leq x_2^*(\mu_1)$ , and  $r_2^*(x,h)=0$  if  $x \geq x_2^*(\mu_0)$ . Note that the above represents extreme cases of retry, i.e., retries of duration zero or infinite.

For the non-extreme case, i.e., the case of  $x_2^*(\mu_1) < x < x_2^*(\mu_0)$ , let  $h^* = \sup\{h \mid r_2^*(x,h)=0\}$ . Since  $r_2^*(x,1)=\infty$  and  $r_2^*(x,0)=0$  for  $x_2^*(\mu_0) < x < x_2^*(\mu_1)$ , we get  $0 \leq h^* < 1$ . For all  $h > h^*$ ,  $r_2^*(x,h) > 0$ , i.e., retry must be applied upon detection of a fault. Suppose retry has been applied for a small duration  $\delta r < r_2^*(x,h)$ . Then, the memoryless property of the exponential distribution leads to the following equation:

$$\rho_3^*(x,h) = (1 - F_1^*(\delta r|h)) (\delta r + \rho_3^*(x, \bar{h}(\delta r))) + \int_0^{\delta r} f_1^*(t|h) (t + \rho_4^*(x,h(t))) dt \tag{18a}$$

By letting  $\delta r \rightarrow 0$  and changing variables, Eq. (18a) becomes

$$\begin{aligned} \frac{d\rho_3^*(x,h)}{dh} &= \left\{ -f_1^*(0|h) (\rho_4^*(x,h(0)) - \rho_3^*(x,h)) \right\} \left\{ \frac{d\bar{h}(r)}{dr} \Big|_{r=0} \right\}^{-1} \\ &= \frac{h\mu_0 + (1-h)\mu_1}{h(1-h)(\mu_0 - \mu_1)} \left\{ \rho_4^*(x,h(0)) - \rho_3^*(x,h) + \frac{1}{h\mu_0 + (1-h)\mu_1} \right\} \end{aligned} \tag{18b}$$

On the other hand,  $\rho_3^*(x,h) = x_0 + t_e$  for all  $h \leq h^*$ . Using the same approach as in Theorem 1, we can prove that  $h^*$  satisfies the following equation

$$\rho_4^*(x,h^*(0)) = x_0 + t_e - \frac{1}{h^*\mu_0 + (1-h^*)\mu_1} \tag{19}$$

From Eq. (4) and the definition of  $\rho_4^*$  in Eq. (13),  $\rho_4^*(x,h)$  is expressed as:

$$\rho_4^*(x,h) = \frac{1}{\nu} (1 - e^{-\nu x}) + e^{-\nu x} \int_0^x \nu e^{\nu y} \rho_3^*(y,h) dy \tag{20}$$

With the initial conditions  $r_2^*(x,1)=\infty$ ,  $\rho_3^*(x,h)$  and  $\rho_4^*(x,h)$  for  $x \leq x_2^*(\mu_1)$ , and Eqs. (18)-

(20), we can calculate  $\rho_k'(x, h)$   $k=3,4$  for all  $x \in (x_2'(\mu_1), x_2'(\mu_0))$  with the following numerical algorithm:

A1. Set  $h=1$ .

A2. Calculate  $\rho_3'(x, 1)$  and  $\rho_4'(x, 1)$  for all  $x \in (x_2'(\mu_1), x_2'(\mu_0))$ .

A3. Calculate  $\frac{d\rho_3'(x, h)}{dh}$  using Eq. (18) and  $\rho_3'(x, h-\delta h)$  for all  $x \in (x_2'(\mu_1), x_2'(\mu_0))$ . (Note  $\rho_3'(x, h)$  and  $\rho_4'(x, h(0))$  are both known.)

A4. Calculate  $\rho_3'(x, h-\delta h)$  using Eq. (20) for all  $x \in (x_2'(\mu_1), x_2'(\mu_0))$ . (Note  $\rho_3'(x, h-\delta h)$  is known for all  $x$ .)

A5. Set  $h=h-\delta h$ . If  $h \leq 0$ , terminate the algorithm.

A6. If  $\rho_4'(x, h(0)) < x_0 + t_s - \frac{1}{h\mu_0 + (1-h)\mu_1}$ , go to A3.

Otherwise, set  $\rho_3'(x, h-\delta h) = x_0 + t_s$  and go to A4.

From the test at A6, one can determine  $h^*$  for all  $x \in (x_2'(\mu_1), x_2'(\mu_0))$  so as to satisfy Eq. (19). Due to the memoryless property of the exponential distribution,  $r_2'(x, h) = 0$  when  $h \leq h^*$  or satisfies Eq. (17) with  $\bar{h}(r) = h^*$  if  $h > h^*$ . In Figure 7,  $r_2'$  versus the prior probability  $h$  is plotted for various values of the residual computation  $x$ . Intersections of the curves in Figure 7 with the horizontal axis give the values of  $h^*$  for different values of  $x$ .

**Remark 1:** in case the active duration of an intermittent fault has a general distribution (instead of an exponential distribution), a differential equation similar to Eq. (18b) cannot be obtained. In such a case, the original integral equation of  $\rho_3(x, \xi_i)$ , i.e., the combination of Eqs. (3) and (12), has to be used instead.

From the foregoing discussion we can determine the optimal retry policy that is based on the prior probability  $h$ . Under this optimal retry policy, we can also determine trajectories of the posterior probabilities after a large number of occurrences and reappearances of intermittent faults have been observed. Let each retry be numbered by a two-tuple  $(m, n_m)$  on the basis of occurrences and reappearances of intermittent faults. The  $(m, n_m)$ -th retry is used to recover from the  $m$ -th occurrence of fault in case of  $n_m=0$  or from the  $n_m$ -th reappearance of the  $m$ -th intermittent fault if  $n_m \neq 0$ . For the hypotheses  $H_i$ ,  $i=0,1$ , let  $h_i(m, n_m)$  represent the posterior probability after the  $(m, n_m)$ -th retry is applied. Also, let  $n_m^t$  be the total number of reappearances of an intermittent fault during the execution of a task and  $h_i(m)$  be the prior probability before the  $m$ -th occurrence of fault, which is equal to  $h_i(m-1, n_{m-1}^t)$  by definition. There are now two main problems to be addressed: (i) Will  $h_i(m)$  converge to either 0 or 1, namely to the true fault characteristic as  $m \rightarrow \infty$ ?; (ii) If converges, how fast will it converge? For convergence, we get the following theorem.

**Theorem 3.** Let  $M = \inf \{m \mid h_i(m) > 1-\epsilon, \text{ or } h_i(m) < \epsilon\}$  where  $0 < \epsilon < 1$ . If  $0 < h_i(0) < 1$ , and  $x_0 + t_0 - \frac{1}{\mu_i} > 0$  for all hypotheses  $H_i$  and all tasks, then  $Prob(M < \infty) = 1$  and  $E[M] < \infty$ .

*Proof:* Let  $S_i(m) = \log \frac{h_i(m)}{h_j(m)}$  for  $j \neq i$ . Thus,  $M$  can be defined as  $\inf \{m \mid |S_i(m)| > K\}$ ,

where  $K = \log(\frac{1-\epsilon}{\epsilon})$ . Let  $z_i(m, n_m) = \log \frac{g[e(m, n_m) | \mu_i]}{g[e(m, n_m) | \mu_j]}$  where  $e(m, n_m)$  is the event

observed at the  $(m, n_m)$ -th retry and  $g(e | \mu_j)$  is the generalized conditional density function defined in Eq. (15). (When the retry duration defined by a retry policy is zero,  $e(m, n_m)$  is null and  $z_i(m, n_m) = 0$ . Also, when  $n_m = 0$ , the retry duration is  $r_1^t$  since the fault type is

not known at its first occurrence.) From Bayes theorem, we have

$$S_i(m') = S_i(m' - 1) + \sum_{n=0}^{n'_m} z_i(m', n) = \sum_{m=1}^{m'} \sum_{n=0}^{n'_m} z_i(m, n) + \log \frac{h_i(0)}{h_j(0)}$$

Let  $y_i(m) = \sum_{n=0}^{n'_m} z_i(m, n)$  under the optimal retry policy.  $S_i(m)$  becomes the sum of independent random variables. After an event is observed, the expected value of  $z_i(m, n_m)$  is the Kullback-Leibler information number and is greater than or equal to zero when  $H_i$  is true [23]. In this case,  $E[z_i(m, n_m)] = 0$  if and only if the prior probability before the  $(m, n_m)$ -th retry is 0 or 1. Since  $x_0 + t_r - \frac{1}{\mu_i} > 0$  for all hypotheses  $H_i$  and all tasks executed,  $Prob(r_i \neq 0) > 0$   $i=1, 2$ . Hence,  $Prob\{y_i(m) = 0\} < 1$  for all  $m < M$ . Following the proof in [24] that the sampling of a sequential probability-ratio test (SPRT) terminates with probability 1,  $Prob\{M < \infty\} = 1$  and  $E\{M\} < \infty$  are obtained. Q.E.D.

**Remark 2:** Since the tasks affected by intermittent faults do not have to be identical, the random variables  $y_i(1), y_i(2), \dots$  are independently but not identically distributed. Moreover, for a fixed  $m$ ,  $z_i(m, n_m)$ 's are dependent on one another because the events observed are controlled by the retry duration which is in turn a function of the moment of reappearance. However, all  $z_i(m, n_m) \geq 0$  when  $H_i$  is true. The condition,  $x_0 + t_r - \frac{1}{\mu_i} > 0$  for all hypotheses  $H_i$  and all tasks executed, indicates that retry is always a useful recovery when an intermittent fault is detected. In fact, this condition is not necessary true for all tasks, but Theorem 3 holds as long as  $Prob(r_i > 0) \neq 0$ .

Theorem 3 shows that the expected number of faults observed -- that makes the posterior probability reach either  $\epsilon$  or  $1-\epsilon$  -- is finite. This also holds for other distributions and retry policies as long as  $r_1 \neq 0$  and  $r_2 \neq 0$  for some  $x$ . However, it does not pro-

vide the average sample size,  $E[M | H_i]$ , that is necessary to reach these termination boundaries  $K$  and  $-K$ . Also, one has to justify whether or not the posterior probability at the termination implies the true fault characteristic. In other words, it is important to know the error probability,  $Prob(S_k(M) < -K | H_i)$ .

There are two difficult aspects in the evaluation of  $E[M | H_i]$  and  $Prob(S_k(M) < -K | H_i)$ ; one is that  $y_i(m)$ 's are not identically distributed, and the other is the non-existence of closed form solutions for both  $r_1^i$  and  $r_2^i$ . If the same task is executed repeatedly under the condition  $x_0 + t_i - \frac{1}{\mu_i} > 0$  for all hypotheses, then  $y_i(m)$ 's become independently and identically distributed. Assume further that initially, both hypotheses are equally likely, i.e.,  $h_0(0) = h_1(0) = 0.5$ . Using the characteristics of SPRT in [20], the error probability is approximated by:

$$Prob(S_k(M) < -K | H_i) \approx \frac{1 - e^{-K}}{e^K - e^{-K}} \approx e^{-K}$$

Even if the same task is executed repeatedly, it is very difficult to obtain an exact solution for  $E[y_i]$  because of the dependency between the optimal retry durations and the observed samples of the active durations. This fact in turn makes it impossible to obtain the exact solution of  $E[M | H_i]$ . Due to the above difficulties, in what follows, we will derive upper and lower bounds of  $E[M | H_i]$  instead of an exact solution.

Suppose there are two retry policies  $R^0$  and  $R^1$  with the retry durations  $(r_1^0, r_2^0)$  and  $(r_1^1, r_2^1)$ , respectively.  $r_1^j(x, h)$  and  $r_2^j(x, h)$  are defined the same as  $r_1^i(x, h)$ .  $r_2^j(x, h)$  is equal to  $\infty$  if  $x \leq x_2^j(\mu_j)$  and 0 otherwise for  $j=0,1$ . Let  $y_i^j(m)$  and  $M^j$  be  $\sum_{n=0}^{n_m} z_i^j(m, n)$  and the number of faults observed to reach the termination boundaries under the retry policy  $R^j$ , respectively. Then, (i)  $Prob(M^j < \infty) = 1$  and  $E[M^j] < \infty$ , and (ii)  $E[y_i^0] \leq E[y_i] \leq E[y_i^1]$ .

(Note that the indices  $m$  are omitted because of the distributions being identical.) Once  $E\{y_j^i | H_i\}$   $j=0,1$  is calculated as in the Appendix, the expected sample size to reach the boundaries  $1-\epsilon$  and  $\epsilon$  is bounded as:

$$E\{M^0 | H_i\} \leq E\{M | H_i\} \leq E\{M^1 | H_i\}$$

where  $E\{M^j | H_i\} \approx \frac{K}{E\{y_j^i | H_i\}}$   $j=0,1$  (see [20] for more on this).

The above equations give the error probability and the bounds of the expected sample size when a certain level of significance is to be achieved. These bounds of  $E\{M | H_i\}$  become tight when the difference between  $\mu_0$  and  $\mu_1$  is small. Of course, the expected sample size under the optimal retry policy is larger than that for the case when the complete information about active duration is observed, i.e.,  $r_1=r_2=\infty$ .

Thus far, we have discussed solutions to the problem of sequential retry decision and hypotheses testing only for the case of exponentially distributed durations. Notice, however, that (i) the same method, with little modification, can be applied to the cases with any other kind of distributions, and (ii) Theorem 3 holds as long as  $Prob\{y(m)=0\} < 1$ . Moreover, the method can be extended to the testing of multiple alternative hypotheses by specifying the prior and posterior probabilities as a vector, each element of which represents the probability that the corresponding hypothesis is true.

## 8. CONCLUSION

In this paper, we have investigated optimal retry policies with *known and unknown* fault characteristics. Retry not only saves the recovery overhead but provides a means to estimate the unknown characteristics. Although the data resulting from retries are censored, they are the only significant means of monitoring the fault characteristics.

Naturally, the monitored results are different from those obtained during device manufacture.

In the discussion of retry policies, retry durations are assumed to be continuous. In fact, the retry durations should be discrete since the time required for repeated execution of an operation cannot be cascaded into a single continuous duration. Since the expected risk is a continuous function of the retry duration, it is not difficult to find the optimal retry policy which is specified as a number of retry attempts.

As was pointed out in the discussion of the expected sample size for reaching a certain level of confidence in hypotheses testing, the test under the optimal retry policy turns out to be inefficient in the sense of maximizing the information observed. This is due to the fact that the optimal retry policy is defined to minimize the total completion time of the task affected by the occurrence of fault. Thus, the retry policy is a local optimum; "optimal" only for the task involved. Clearly, the retry policy which gives complete maximum information should have infinite retry durations, although such a retry policy is totally unacceptable in reality. It would be interesting to examine the trade-off between the two extreme objectives, i.e., minimizing the local task completion time and maximizing the information to be collected. This problem can be formulated as the minimization of the asymptotically accumulated risk,  $\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{j=1}^m E[\rho_k(x, C_j^i)]$ , where  $j$  and  $m$  are used to number the successive retries and  $C_j^i$  is the measured fault characteristic at the  $j$ -th retry. It can also indicate that the global optimal retry policy should collect more information (it is definitely not complete though) from the beginning to speed up the estimation of the true fault characteristics and then implement the local optimal retry policy once the true characteristics are obtained.

Another important aspect is the choice of an accurate model for the fault behavior. As was discussed in Sections 4 and 5, the optimal retry policy and the measurement of the fault characteristics are dependent on the family of density functions that are initially selected. The suitability of chosen models can be validated through goodness-of-fit tests, e.g., chi-square goodness-of-fit. Although sometimes the expected task completion time may not be minimized because of the poor choice of model, the information collected via retries can still be used to check the suitability of the model. Thus, after a large number of samples have been obtained, it is possible to select an appropriate form of density function and then achieve the minimum task completion time. The other approach is to begin with hypotheses of various forms of density functions. As sampling progresses, the parameters associated with the density function forms are estimated and then the hypotheses are tested.

The work presented in this paper is to incorporate the capability of real-time estimation (of the fault characteristics) and decision (on optimal retry policies) into the computer system. The results are a self-adjustable (thus intelligent) system and a powerful measurement of the fault characteristics. This idea can also be extended to other applications, e.g. the measurement of program behavior and the simultaneous decision of system configuration or scheduling. Such extensions would be significant contributions towards the construction of highly intelligent computer systems.

#### ACKNOWLEDGMENT

The authors are grateful to C. M. Krishna at the University of Michigan for his careful comments on the first draft of this paper, and to both R. Butler and M. Holt at

NASA Langley Research Center for their technical and financial assistance.

## REFERENCES

- [1] D. P. Siewiorek and R. S. Swarz, *The Theory and Practice of Reliable System Design*, Digital Press, 1982.
- [2] O. Tasar and V. Tasar, "A Study of Intermittent Faults in Digital Computers," *1977 AFIPS Conference Proceedings*, Vol. 46, 1977, pp. 807-811.
- [3] M. Ball and F. Hardie, "Effects and Detection of Intermittent Failures in Digital Systems," *1969 AFIPS Conference Proceedings*, Vol. 35, Fall 1969, pp. 329-335.
- [4] K. G. Shin and Y. H. Lee, "Error Detection Process: Model, Design, and Its Impact on Computer Performance," To appear in *IEEE Trans. on Computers*, Vol. C-33, No. 6, June 1984.
- [5] W. C. Carter, et al., "Cost Effectiveness of Self Checking Computer Design," *Proc. of the 7th Annual Int'l Symp. on Fault-Tolerant Computing*, 1977, pp. 117-123.
- [6] G. H. Maestri, "The Retryable Processor," *1972 AFIPS Conference Proceedings*, Vol. 41, Fall 1972, pp. 273-277.
- [7] D. L. Droulette, "Recovery through Programming System/300-System/370," *1971 AFIPS Conference Proceedings*, Vol. 38, Spring 1971, pp. 467-476.
- [8] L. A. Boone, H. L. Liebergot and R. M. Sedmak, "Availability, Reliability, and Maintainability Aspects of the SPERRY UNIVAC 1100/60," *Proc. of the 10th Annual Int'l Symp. on Fault-Tolerant Computing*, 1980, pp. 3-9.
- [9] J. J. Stiffler and L. A. Bryant, "CARE III Phase Report - Mathematical Description," *NASA Report*, No. 3566, Nov. 1982.
- [10] I. Koren and S. Y. H. Su, "Reliability Analysis of N-Modular Redundancy Systems with Intermittent and Permanent Faults," *IEEE Trans. on Computers*, Vol. 28, No. 7, July 1979, pp. 514-520.
- [11] E. Cinlar, *Introduction to Stochastic Processes*, Prentice-Hill Inc., 1975.

- [12] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, 1983.
- [13] A. C. Cohen, "Progressively Censored Samples in Life Testing," *Technometrics*, Vol. 5, No. 3, Aug. 1963, pp. 327-339.
- [14] A. C. Cohen, "Multi-censored Sampling in the Three Parameter Weibull Distribution," *Technometrics*, Vol. 17, No. 3, Aug. 1975, pp. 347-351.
- [15] A. C. Cohen, "Progressively Censored Sampling in the Three-Parameter Gamma Distribution," *Technometrics*, Vol. 19, No. 3, Aug. 1977, pp. 333-340.
- [16] G. H. Lemon, "Maximum Likelihood Estimation for the Three Parameter Weibull Distribution Based on Censored Samples," *Technometrics*, Vol. 17, No. 2, 1975, pp. 247-254.
- [17] D. R. Wingo, "Solution of the Three-Parameter Weibull Equations by Constrained Modified Quasilinearization (Progressively Censored Samples)," *IEEE Trans. on Reliability*, Vol. R-22, No. 2, June 1973, pp. 96-102.
- [18] E. L. Lehmann, *Testing Statistical Hypotheses*, John Wiley & Sons, 1959.
- [19] W. R. van Zwet, "Bias in Estimation from Type I Censored Samples," *Statist. Neerlandica*, Vol. 20, 1966, pp. 143-148.
- [20] M. H. DeGroot, *Optimal Statistical Decision*, McGraw-Hill Book Company, 1970.
- [21] A. Irle and N. Schmitz, "Decision Theory for Continuous Observations I: Bayes solutions," *Trans. of the 7th Prague Conf. on Information Theory, Statistical Decision Functions and Random Processes*, 1974, pp. 209-221.
- [22] I. O. Berger, *Statistical Decision Theory*, Springer-Verlag, New York, 1980.
- [23] H. Chernoff, *Sequential Analysis and Optimal Design*, S.I.A.M., 1972.
- [24] C. Stein, "A note on cumulative sums," *Ann. Math. Stat.*, Vol. 17, 1946, pp. 489-499.

**Appendix: The Expression of  $E\{y_1^j|H_i\}$**

The retry duration  $r_2^j$  under the retry policy  $R^j$  is equal to  $\infty$  if  $x \leq x_2^j(\mu_j)$  and 0 otherwise. Thus, the complete information will be gathered if an old intermittent fault is detected again at  $x \leq x_2^j(\mu_j)$  and no information will be obtained if detected at  $x > x_2^j(\mu_j)$ . Hence, if the retry for a newly detected intermittent fault when the residual computation is  $x$  succeeds, we expect to collect information from the successive retries before the task completion as follows:

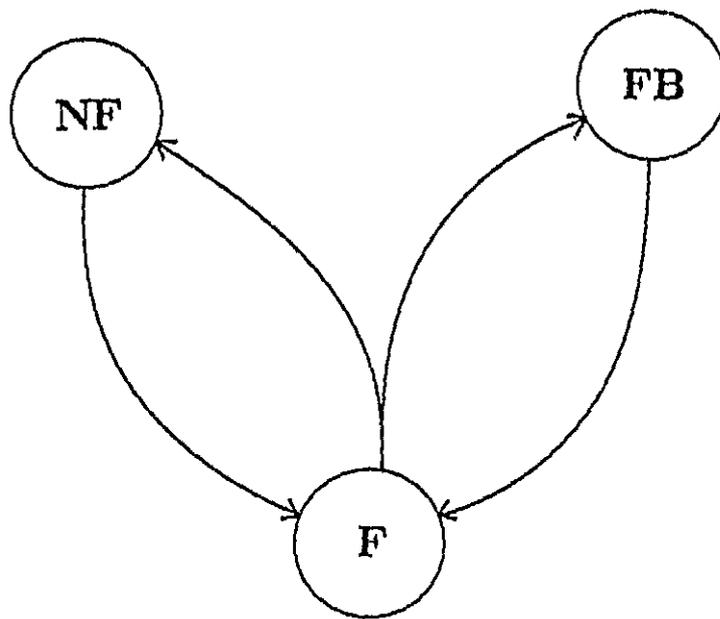
$$E\{I_2|x\} = E\left\{\sum_{n=1}^{n'_m} z_i(m,n)|x\right\} = \begin{cases} \nu x \left(\log \frac{\mu_i}{\mu_j} - \frac{\mu_i - \mu_j}{\mu_i}\right) & \text{if } x \leq x_2^j(\mu_j) \\ e^{-\nu(x-x_2^j(\mu_j))} \nu x_2^j(\mu_j) \left(\log \frac{\mu_i}{\mu_j} - \frac{\mu_i - \mu_j}{\mu_i}\right) & \text{otherwise} \end{cases}$$

Let the maximum retry duration for a newly detected fault be  $r_1^j(x)$  when the residual computation is  $x$ . Also, let  $\psi(x_d)$  be the density function of the detection time of a new intermittent fault,  $x_d$ , given that it is detected during the task execution. Then,

$\psi(x_d) = \frac{\lambda_i e^{-\lambda_i x_d}}{1 - e^{-\lambda_i x_0}}$  where  $\lambda_i = p_i \lambda$ . Thus, we have  $E\{y_1^j|H_i\}$  as follows:

$$E\{y_1^j|H_i\} = \int_0^{x_0} \psi(x_0-x) e^{-\mu_i r_1^j(x)} I_1^j(r_1^j(x)) dx + \int_0^{x_0 r_1^j(x)} \int_0^t \mu_i e^{-\mu_i t} \psi(x_0-x) \{I_1^j(t) + E\{I_2|x\}\} dt dx$$

where  $I_1^j(r) = -(\mu_i - \mu_j)r$  is the information collected from an unsuccessful retry of the maximum retry duration  $r$  and  $I_1^j(r) = \log \frac{\mu_i}{\mu_j} - \frac{\mu_i - \mu_j}{\mu_i} r$  is the resulting information when the retry succeeds after the duration  $r$ .



NF: non-faulty  
F: faulty  
FB: fault-benign

Figure 1. The Model of Faults.

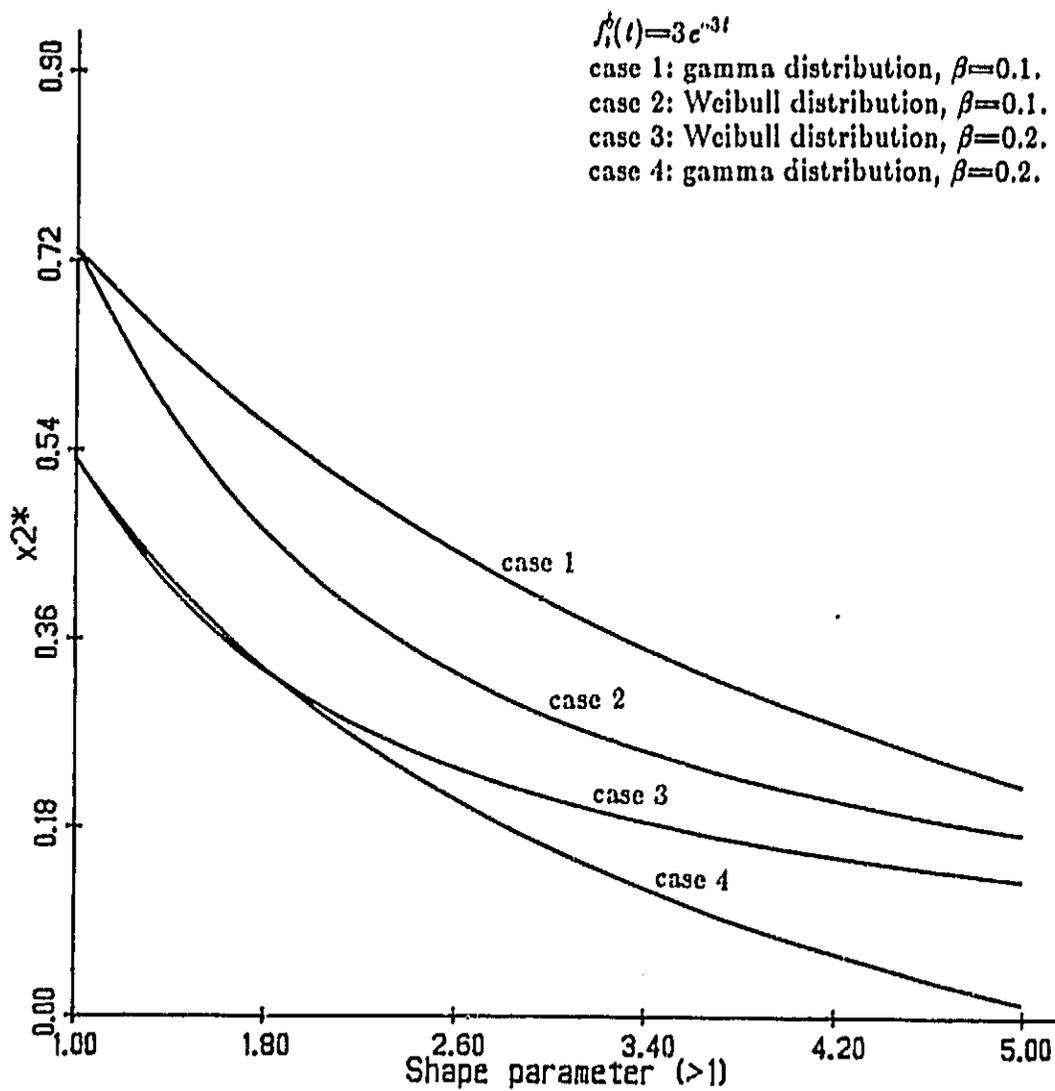
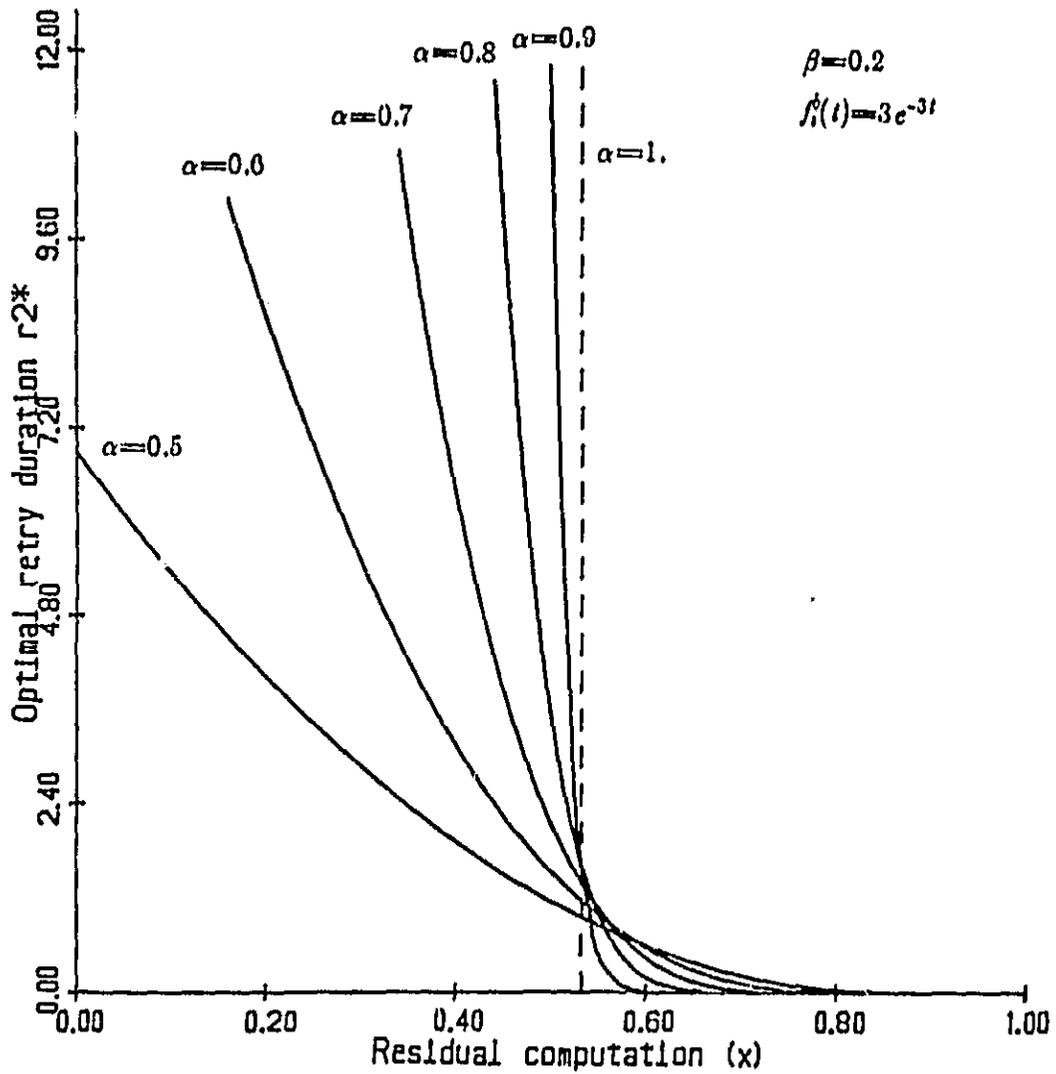
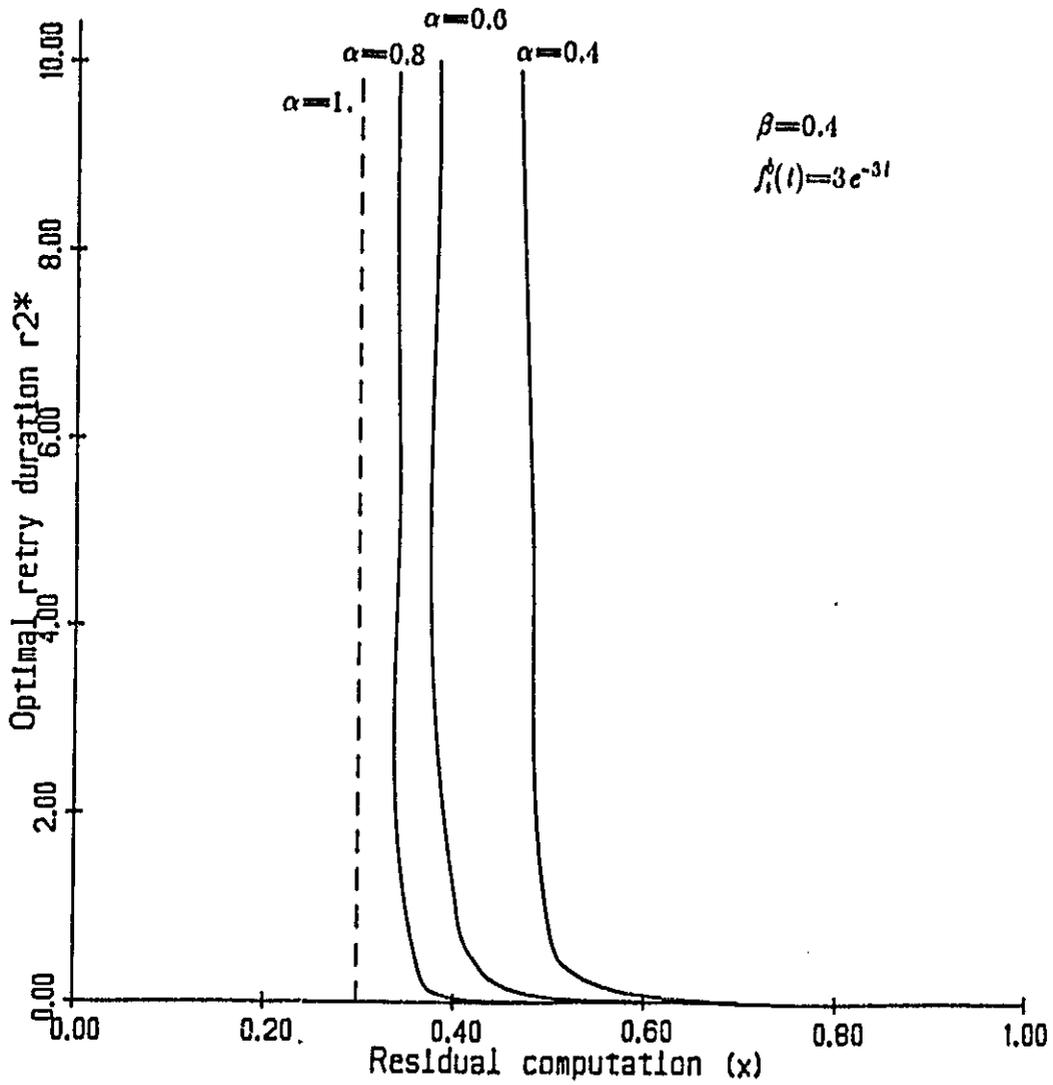


Figure 2.  $x_2^*(\alpha, C_f)$  versus the Shape Parameter  $\alpha$  when the Hazard Rate is Increasing.



The density function of active duration  $f_i(t) = \frac{\alpha}{\beta} t^{\alpha-1} e^{-\frac{t^\alpha}{\beta}}$

Figure 3. The Optimal Retry Duration  $r_2^*(x, C_j)$  for Weibull Distributions with Decreasing Hazard Rate.



The density function of active duration  $f_i^0(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}}$

Figure 4. The Optimal Retry Duration  $r_2^*(x, C)$  for Gamma Distributions with Decreasing Hazard Rate.

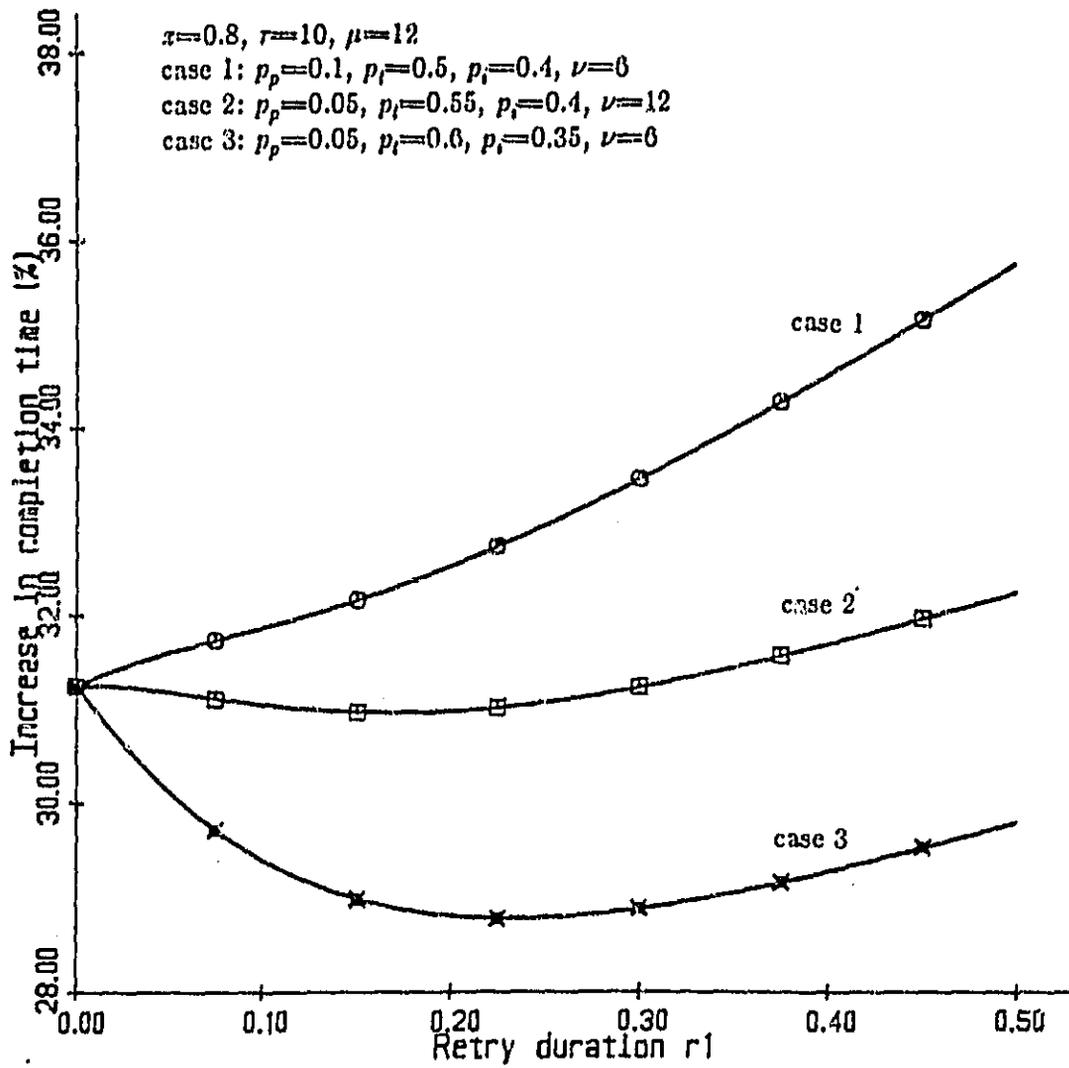


Figure 5.  $V_2(x, C, R) - x$  versus the Retry Duration  $r_1$  for Exponentially Distributed Durations.

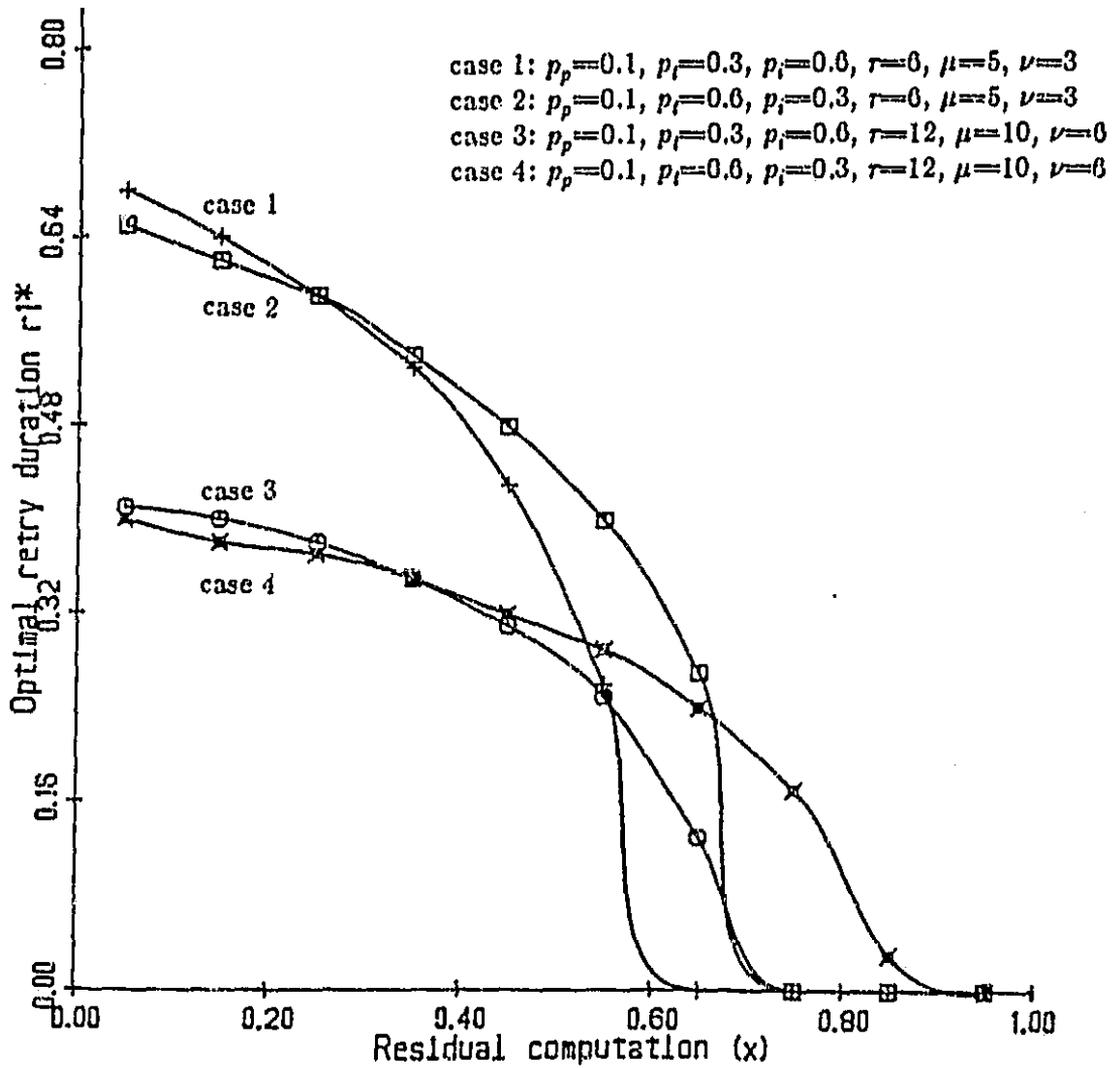


Figure 6. The Optimal Retry Duration  $r_1^*(x, C_j)$  for Exponentially Distributed Durations.

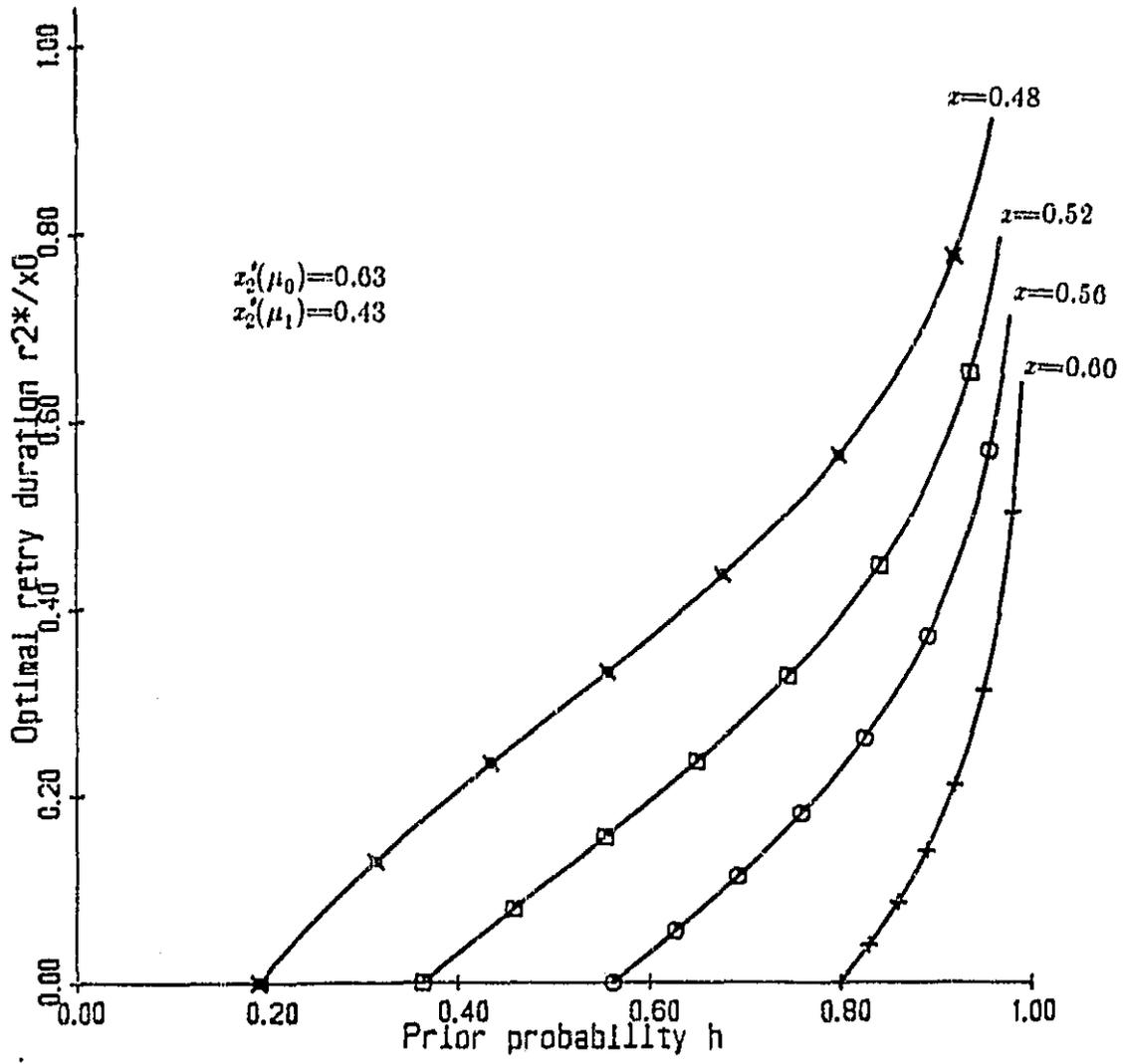


Figure 7. The Optimal Retry Duration  $r_2^*(z, h)$  versus the Prior Probability  $h$ . ( $\mu_0=10, \mu_1=5, \nu=5$ )