Reliability of Fault Tolerant Control Systems: Part I

N. Eva Wu
Department of Electrical Engineering, Binghamton University
Binghamton, NY 13902-6000, U.S.A.
Tel: 607-777-4375, Fax: 607-777-4464, Email: evawu@binghamton.edu

Abstract

This paper reports Part I of a two part effort that is intended to delineate the relationship between reliability and fault tolerant control in a quantitative manner. Reliability analysis of fault-tolerant control systems is performed using Markov models. Reliability properties peculiar to fault-tolerant control systems are emphasized. As a consequence, coverage of failures through redundancy management can be severely limited. It is shown that in the early life of a system composed of highly reliable subsystems, the reliability of the overall system is affine with respect to coverage, and inadequate coverage induces dominant single point failures. The utility of some existing software tools for assessing the reliability of fault tolerant control systems is also discussed. Coverage modeling is attempted in Part II in a way that captures its dependence on the control performance and on the diagnostic resolution.

1 Introduction

Highly reliable systems make use of redundancy to achieve fault tolerance, due to limited reliability of components or subsystems [4]. Utilization of analytic redundancy [5] that provided by static and dynamic relations among system variables, such as secondary functions of effectors, virtual measurements, projections, etc. can further reduce the probability of exhaustion of hardware in a cost-effective manner. Analytic redundancy management of complex control systems, however, involves considerable more risks in comparison with such schemes as majority voting, for decision making is often based on residual signals formed by the differences between noisy measurements and calculated values of output variables based on inaccurate models. Decision errors can be associated with uncertainties on whether there is a subsystem failure, which subsystem has failed, how severe is its effect, whether it is necessary to take a drastic corrective action, which action to take. In addition, the question may also arise on whether there is adequate control relevant redundancy [15] and authority to allow recovery from the effect the failure. The dynamic and closed-loop nature, common to all control systems, is the source for additional difficulties, such as temporary mask of the effect of subsystem failures, the vagueness in the definition of a system level failure in the context of control performance, and the sometimes significant processing requirement in supporting the redundancy management.

Definitions suggested in [9] on fault and failure are adopted with a slight extension. A fault is an unpermitted deviation of at least one characteristic property or variable of the system. A failure is a permanent interruption of a system's ability to perform a required function under specified operating conditions. Note that a failure can also be defined in the subsystem level. A failure may or may not lead to a failure. Without loss of generality, a subsystem failure is assumed to always lead to the system failure unless a successful management of redundancy ensues. A system level failure is declared when faults or subsystem failures cause the control performance of the system to fall below the prescribed threshold. The performance threshold can be set at two (or more) different levels, each corresponding to a specific reliability requirement. In aviation, for example, one level can be set by the ability to carry out a normal mission (or mission abort in terms of failure probability), and another can be set by the ability to merely maintain the system stability needed for safe landing (loss of control in terms of failure probability). This paper will treat different reliability requirements in a unified manner.

Reliability is naturally a subjective concern in the analysis and design of fault-tolerant control systems. Few publications that formally address this issue [10, 11, 13] have confined the scope of discussion to reliability assessment for dynamic systems subject to faults. Reliability is rarely regarded as an objective
criterion that guides a control system design in an integrated manner. This predicament is due to the difficulty in establishing a functional linkage between the overall system reliability, and the performance defined in the conventional sense at the bottom level for controls and for diagnosis. The only attempt prior to this work along this direction is reported in [11] where such a linkage is established through coverage under the possibilistic formalization. The possibilistic formalization provides flexible and usually more accurate descriptions of uncertainties, but suffers from lack of corresponding theoretical and numerical means for reliability analysis. This paper is intended to address the reliability issue of fault tolerant control systems in the more familiar probabilistic formalization so that existing tools and methodologies of reliability analysis can be applied. The paper is organized as follows. Section 2 presents issues encountered in our endeavor through a reliability analysis case study of a fault tolerant flight control system. Section 3 discusses numerical techniques for reliability assessment of fault tolerant systems with emphasis on how coverage enters reliability as a decision risk factor. Several approximate relations are derived to reveal the dependence of reliability to coverage in simple forms.

2 A Case Study

To understand some of the reliability issues peculiar to fault tolerant control systems, we start with a reliability assessment case study of a fault tolerant flight control system (FTFCS). A complete report on this case can be found in [13]. Observations that follow should serve to motivate a focused effort in coverage modeling of fault tolerant control systems.

![Effector functional dependency in a FTFCS](image)

Fig.1 Effector functional dependency in a FTFCS

Fig.1 shows the functional dependencies of subsystems in the pitch/roll control effector block within a fault tolerant flight control system [13]. The diagram reflects the available redundant lateral control author-

The reliability indicator used in the following discussion is the probability of loss of control denoted by $P_{LOC}$. Each small box in Fig.1 represents a subsystem, where $\lambda_j(X = I, A, S)$ are the failure rates in terms of failures per hour. Under the assumption of low subsystem failure rates, short mission time, and highly rigorous maintenance requirements, constant failure rates are appropriate. Safety requirement for inner layer parallel configuration (the n-plex computer/effector interface subsystem) considered is 1-out-of-n. Safety requirement for the outer layer parallel configuration is 3-out-of-4. This means that the three remaining effector channels must work in concert to accommodate a failure in one effector channel.

The redundancy architecture shown in Fig.1 does not truly reflect how effector channel hardware is configured. It must be understood as an effective redundancy configuration which assumes that any anomaly in an effector channel serious enough to warrant a control adaptation or reconfiguration action for failure accommodation can do so promptly and successfully. In reality, however, due to uncertainties in the model of the system to be controlled, uncertainties in the models of signals exerted on the system, and the limited processing capability, considerable risks exist in making a decision on a corrective action. These decision risks must be taken into consideration in reliability assessment. The risks encountered may include overly slow or severe transients, false alarm, miss detection, false identification, false reconfiguration, and lack or exhaust on of redundancy. The notion of coverage is now used to account for such risks. It represents an attempt to separate handling of failures from occurrence of failures. Coverage defined in this context is highly scenario dependent, highly time dependent, and most of all, difficult to estimate. Coverage has been used as a parameter to reflect the ability of a system to automatically recover from the occurrence of a fault during.
a normal system operation[6].

Coverage \equiv Probability(System recovers|Fault occurs).

Once a decision is made however, the process of removing a subsystem or reconfiguring the system is generally involved. This process, though fast in comparison with a failure process, still takes time, and has been shown to be generally non-exponentially distributed. Including this process in a reliability model implies the creation of a numerically stiff problem[4].

Some results of reliability assessment for the system of Fig.1 are now presented. All coverage values are obtained based on test data[16], which aggregate the effects of decision errors. Since these values are fixed, they are called static coverage values. A coverage value of 0.99 is used when an actuator failure is accommodated. The following table gives coverage of a computer/effector interface subsystem failure.

<table>
<thead>
<tr>
<th>Redundancy management</th>
<th>Number of intact subsystems</th>
<th>Coverage ( c_{4-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority voting</td>
<td>4</td>
<td>0.992</td>
</tr>
<tr>
<td>Majority Voting</td>
<td>3</td>
<td>0.99</td>
</tr>
<tr>
<td>Comparing</td>
<td>2</td>
<td>0.89</td>
</tr>
<tr>
<td>Self-monitoring</td>
<td>1</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 1

Coverage associated with surface damage is left as a variable whose required value is yet to be determined for the reason that it is where improvement is needed most. A realistic estimate of static coverage can be obtained by counting the number of unsuccessful surface failure recoveries and taking the ratio with respect to the total number of simulated surface impairments with a full scale simulator. Note that such static coverage values infer from a rather small sample of coverage data to a general population, which do not address a specific process well, and therefore are inadequate for use to make online decisions. Section 4 will discuss coverage modeling for more accurate coverage prediction.

The approximate parameter ranges in the Markov used in our case study are now given. The overall system reliability is required to achieve \( 1 - 10^{-7} \).

Subsystem failure rate \( \lambda_X \) \( 10^{-6} \sim 10^{-4} \) hour\(^{-1}\)

Subsystem mean time to recover \( \mu_j \) \( 10^{-3} \sim 10^{-4} \) hours

Variance of time to recover \( \sigma_j \) \( 10^{-3} \sim 10^{-4} \) hours

Mission time \( T \) \( 10^3 \sim 10^4 \) hours

Table 2

The above table reflects two common characteristics of highly reliable fault tolerant systems: details due to small failure probabilities cannot be arbitrarily ignored, and recovery process is much faster than failure process (10\(^7\) times faster at least). As a result, one is faced with solving a numerically stiff problem. Fortunately, successful attempts have been made to effectively deal with the stiff problem both theoretically and numerically[12, 4].

Under a set of given failure rates and mission time, the following results are obtained for the effector block coverage associated with surface damage. These claims will be affirmed through analytical means in the next section. The potential benefit of enhancing coverage and the potential cost of additional hardware redundancy have now given us sufficient motivation to investigate what factors affect coverage and in what ways coverage is affected in a fault tolerant control system.

### 3 Coverage in Reliability Assessment

In this section, the development of reliability model and the numerical technique used for obtaining the results of the previous section are presented. Several general results regarding the critical role of coverage in reliability assessment are then derived. Coverage modeling, calculation, and its role in relating fault tolerant control to reliability will be discussed in the next section.

Reliability modeling can be regarded as a process of identifying the structure function of a system comprised of \( N \) subsystems with positive random lifetimes. The structure function defines a mapping: \( \{0, 1\}^N \rightarrow \{0, 1\} \). Reliability assessment can be regarded as a process of evaluating the mapping, given state transition probabilities. A subsystem is in state “1” (intact) before its lifetime and state “0” (failed) after its lifetime. The fundamental assumption of a Markov process is that the probability that a system will undergo a transition from one state to another state depends only on the current state of the system and not any previous states the system may have experienced. A Markov process where all state transition rates are time-invariant is said to be homogeneous.

Keeping in mind the case study of the previous section, the following assumptions are used in the subsequent development of reliability models:

(a) all subsystems are operational at the onset;
(b) failure probability of any given subsystem is \( 1 - e^{-\lambda} \) where \( \lambda \) is the constant failure rate of that subsystem;
(c) a failure in any subsystem is independent of that in
all other subsystem failures;
(d) redundancy management restores the system operation with a certain coverage following a subsystem failure;
(e) an uncovered subsystem failure always leads to the system failure (caused by decision errors, delays in redundancy management, and the exhaustion of redundancy as a special case);
(f) a covered subsystem failure obeys a recovery time distribution with mean time $\mu$ and variance $\sigma^2$ (caused by transients following the removal of a failed subsystem, or reconfiguration of control law);
(g) all rates of recovery are orders of magnitude faster than rates of subsystem failures;
(h) a failed subsystem remains failed during a mission (no repair).

In the following development, the parameter values given in Table 2 are assumed. The calibration to a particular time scale allows us to draw more useful conclusions for similar classes of applications. With the framework set, we now demonstrate the role of coverage through a progressively more complex reliability model starting from the inner most layer of the two-layer parallel-to-series interconnection scheme shown in Fig.1. The rate diagram shown in Fig.2 represents a system comprised of four computer-effector interface subsystems with subscript "I" omitted for simplicity. This represents a structure of a well-studied $k(1)$-out-of-$n(4)$ system (any combination of $k$ operating subsystems out of $n$ independent subsystems will guarantee successful operation of the system [7]). The complication here comes with the time varying recovery rates and with the incomplete reconfiguration. Our main interest lies with the role of coverage in the context of fault tolerant control, and derive some relations that in Table 1 will not be needed for the moment, but will enter the picture, when a second parallel layer is in place. In Fig.2, states are denoted by circled two-digit numbers. State name $ij$, $i > 0$ indicates that there are $i$ intact subsystems and $j$ failed subsystems. Subscript $0j$ always denotes an exit (death) state. The coefficient in front of the transition rate $\lambda_j(t)$ represents the number of subsystems that can fail in that transition. A transition to a recovered state is marked by transition rate $\lambda_j(t)$, which, in terms of conditional transition time distribution function $f_j(t)$, is given by

$$\lambda_j(t) = \frac{f_j(t)}{1 - F_j(t)}, \quad F_j(t) = \int_0^t f_j(\tau)d\tau$$

where $j$ is the total number of failed subsystems at the time of the transition. Denote the mean time and variance of this transition by $\mu_j$ and $\sigma_j$, respectively. Coverage associated with the transition out of the state with $j$ failed subsystems is denoted by $c_j$. The values of $c_j$ used for the system shown in Fig.2 are given in Table 2 of the previous section. Aggregating all exit states in one with failure probability $p_D(t)$, the state-space dimension of the Markov model is reduced to eight. From the rate diagram, a system of ordinary differential equations

$$\dot{P}(t) = P(t)Q(t), \quad P(0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

can be derived [8], where

$$P(t) = \begin{bmatrix} p_{00}(t) & p_{01}(t) & p_{02}(t) & p_{03}(t) & p_{10}(t) & p_{20}(t) & p_{11}(t) & p_{D}(t) \end{bmatrix},$$

is the vector of probabilities of holding at state $ij$ at time $t$, and $Q(t)$ is the state transition rate matrix,

Let $\Phi(t, T)$ denote the solution to $\Phi(t, T) = P(t)\Phi(t, t)$, $\Phi(0, T) = I$. The system failure probability at the end of mission time $T$ is given by $p_D(T) = P(T)\Phi(t, T)$.
Given the large disparity between failure rates ($\lambda$) and recovery rates ($1/\mu_i$) as shown in Table 2, it is meaningful to examine the condition under which the recovery times in the Markov model can be eliminated. The rationale for this intent lies with the simplification to a homogeneous Markov process. Suppose this elimination is allowed, the Markov model will have been simplified to that depicted in Fig. 3(a). The sum of the probabilities of the death states that have been eliminated as a result of ignoring the recovery time is now estimated. First a result on an approximate failure probability is given.

**Theorem 1.** Assume (a) through (h) hold for a $k$-out-of-$n$ system. In addition, $c_0 < 1$, and $n\lambda T << 1$. Then the system failure probability is dominated by

$$P^A_D(T) = n\lambda T (1 - c_0)$$

if

$$1 - c_0 >> \frac{(n-1)\lambda T (1 + \lambda T)^n - 1 + (1 + \lambda T)^n - (1 + n\lambda T)}{n\lambda T (1 - \frac{1}{\mu_i})}$$

where

$$\mu = \max\{\mu_1, \mu_2, \ldots, \mu_{n-k}\}, \mu_i = 0, \forall i > n-k$$

In this case the approximation error $|P_D(T) - P^A_D(T)|$ satisfies

$$\frac{|P_D(T) - P^A_D(T)|}{P_D(T)} < \max\{(n-1)\lambda T (1 + \lambda T)^n - 1, -(1 + \lambda T^n) - (1 - n\lambda T), \frac{(n\lambda T)^2}{\mu_i} (1 - c_0)\}$$

A key step to proving Theorem 1 is the application of White's bounds[4] by which it is required that state transitions be considered as separate transitions representing disjoint events of traversing paths to exit states. The rest of the proof involves employing adjusting the bounds using Binomial forms. Due to paper length limitations, proofs for all theorems are omitted.

Essentially, Theorem 1 states that if $c_0$ is not sufficiently close to 1 in the sense defined above, the failure probability of the Markov process becomes linear with respect to $\lambda T$, and to $1 - c_0$. The most important implication here is that in order to effectively take advantage of redundancy, it is crucial to have the highest possible coverage for the first failure that occurs in the system. As will be shown in the next section, this can be achieved only through integrated design of the entire system. To gain a sense on how far $c_0$ must be from 1 in order for the simple formula to be valid, values given in Table 2 are used. With $n = 4$, $k = 1$, $T = 1$, $\mu = 10^{-3}$ and $\lambda = 10^{-4}$

$$1 - c_0 >> \frac{3\lambda T (1 + \lambda T)^4 - 1 + (1 + \lambda T)^4 - (1 + 4\lambda T)}{4\lambda T (1 - 2\lambda T)} = 0.00018$$

must be satisfied. Using $c_0 = 0.992$ from Table 1, $1 - c_0 = 1 - 0.992 = 0.008 >> 0.00018$. The following approximation on the system failure has been obtained

$$P^A_D(T) = 3.2 \times 10^{-5}$$

with an approximation error bounded by $6.0 \times 10^{-8}$.

The inequality and error bound in Theorem 1 becomes more and more conservative as $k$ becomes larger and larger than 1, for more terms are added without being subtracted in completing the binomial forms. When $c_i$'s are closer to 1, tighter bounds can be obtained by considering $k\lambda c_i$ as a fast rate relative to $k\lambda (1 - c_i)$.

The next result states the condition on the elimination of the recovery times. This is equivalent to setting $\mu_i = 0, \forall i$.

**Theorem 2.** Assume (a) through (h) hold for a $k$-out-of-$n$ system. In addition, $c_0 < 1$, $n\lambda T << 1$. In addition, assume $p^A_D(T) = n\lambda T (1 - c_0)$ dominates $P_D(T)$. Then the recovery times can be ignored in the system failure probability calculation with an error $e_0$ bounded by

$$n\lambda \mu_i (1 + \lambda T)^n - 1, \mu_i = \max\{\mu_1, \mu_2, \ldots, \mu_{n-k}\}$$

where $\mu_i = 0, \forall i > n-k$. In this case,

$$1 - c_0 >> \frac{(n-1)\mu (1 + \lambda T)^n - 1 + (1 + \lambda T)^n - (1 - n\lambda T)}{n\lambda T (1 - \frac{1}{\mu_i})}$$

$$p_D(T) = \sum_{n+k}^{\infty} \binom{n+k}{k} \left(1 - \frac{1}{\mu_i}\right)^k \left(\frac{1}{n\lambda T (1 - \frac{1}{\mu_i})}\right)^n$$

The proof of Theorem 2 is similar to that of Theorem 1. In fact the above inequality is implied by (4). To see how easily this condition is satisfied, the right hand side is calculated using again $n = 4$, $k = 1$, $\mu = 0.001$, $T = 1$, and $\lambda = 10^{-4}$, which leads to $1 - c_0 >> 3 \times 10^{-7}$. This is met if $1 - c_0 = 10^{-5}$, or $c_0 = 0.99999$. In comparison with the used value of $c_0 = 0.992$, the condition for eliminating recovery time is well satisfied. Therefore, whenever the failure probability is dominated by $n\lambda T (1 - c_0)$, elimination of recovery time is permissible. In the case of the system of Fig. 2, $p_D(T)$ has can be now be expressed analytically as

$$p_D(T) = \sum_{n+k}^{\infty} \binom{n+k}{k} \left(1 - \frac{1}{\mu_i}\right)^k \left(\frac{1}{n\lambda T (1 - \frac{1}{\mu_i})}\right)^n$$

We now proceed to demonstrate the solution procedure for the more complex two-layer parallel-to-series interconnection scheme encountered in the case study of Section 2. Since the achievable $1 - c_0$ in our case study satisfies the condition of (4), the effector block failure probability can be approximated by, after the application of (3) to both the inner and the outer layers of parallel interconnections

$$p_D(T) = 4\lambda T (1 - c_0) + \lambda_1 T (1 - c_0) + \lambda_2 T (1 - c_0)$$. (9)

where $n = 1, 2, 3$, or 4 is the redundancy level in the computer/effector interface portion. This formula holds when $1 - c_0^X > \lambda X T$. $X = I, A, S$. In particular, improvement in coverage, even by a small percentage (from .99 to .999, for example) could reduce the system failure probability by an order of magnitude. On the other hand, the level of redundancy in the interface portion in each effector channel deserves the consideration for optimization. The following table summarizes the contribution of the first term in (9) (the interface portion) to system failure probability.
The numbers in the last column are the products of $n$ and $1 - c'_0$ for different redundancy level $n$. Since $c'_0$ is a decreasing function of $n$ as shown in the table, it turns out that the minimum appears at $n = 3$, i.e., the 3-plex interface architecture minimizes the system failure probability.

Enhanced coverage has been shown to be the key to enhanced system reliability. There are applications, such as civil aviation, where system reliability requirement is as stringent as, for example, $1 - 10^{-9}$. Given the limitation of individual subsystem reliability, the analysis of this section concludes that coverage of first subsystem failures in such systems must be raised to a value extremely close to 1 to avoid inducing dominant single point failures. At this extremely high coverage, the approximate formulae given in this section is no longer accurate, and the use of an elaborate and rigorous numerical tool such as WinSURE becomes necessary. (For the data given in Table 2, however, the failure probability calculation result of the above approximate formula is indistinguishable from the upper and the lower bounds given by ASSISTciteASSIST and SUREciteSURE.) High coverage, at the same time, imposes extremely stringent requirement on redundancy management. Such a requirement must be reflected at the bottom levels on the control and diagnostic performance requirements, which will be discussed in part II of the paper.

<table>
<thead>
<tr>
<th>Redundancy Management</th>
<th>Coverage $c'_0$</th>
<th>Effect on $P_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Majority voting 0.992</td>
<td>$4\lambda T \times 0.032$</td>
</tr>
<tr>
<td>3</td>
<td>Majority Voting 0.99</td>
<td>$4\lambda T \times 0.030$</td>
</tr>
<tr>
<td>2</td>
<td>Comparing 0.89</td>
<td>$4\lambda T \times 0.220$</td>
</tr>
<tr>
<td>1</td>
<td>Self-monitoring 0.75</td>
<td>$4\lambda T \times 0.25$</td>
</tr>
</tbody>
</table>

Table 4

Enhanced system reliability has been shown to be the key to enhanced system reliability. There are applications, such as civil aviation, where system reliability requirement is as stringent as, for example, $1 - 10^{-9}$. Given the limitation of individual subsystem reliability, the analysis of this section concludes that coverage of first subsystem failures in such systems must be raised to a value extremely close to 1 to avoid inducing dominant single point failures. At this extremely high coverage, the approximate formulae given in this section is no longer accurate, and the use of an elaborate and rigorous numerical tool such as WinSURE becomes necessary. (For the data given in Table 2, however, the failure probability calculation result of the above approximate formula is indistinguishable from the upper and the lower bounds given by ASSISTciteASSIST and SUREciteSURE.) High coverage, at the same time, imposes extremely stringent requirement on redundancy management. Such a requirement must be reflected at the bottom levels on the control and diagnostic performance requirements, which will be discussed in part II of the paper.

4 Conclusions

The main contributions of the paper are presented in Theorems 1 and 2.

Theorem 1 states that when coverage is not sufficiently high, the uncovered subsystem failures dominate the system failure, and the system failure probability increases linearly with decreasing coverage values. This can significantly undermine the benefit of using redundancy. Therefore, every effort should be made to enhance coverage of first subsystem failures. Theorem 2 states that when the uncovered failures are dominating, the recovery times can be ignored if they are several orders of magnitude faster than the subsystem failure times on average. In this case, a numerically stiff problem is avoided, and reliability analysis of a complex system can be much simplified.

It is necessary to point out that the motivating force of this work comes from the set goals of the on going NASA/IAA aviation safety program[2]. Though the main conclusions drawn in this paper should hold for many areas of applications, the reader is cautioned to pay attention to the conditions stated upon which the conclusions are drawn, especially when they are employed to applications of vastly different time scales.

References