Proceedings of the 1989 Winter Simulation Conference
E.A. MacNair, K.J. Musselman, P. Heidelberger (eds.)

MONTE CARLO SIMULATION OF COMPLEX SYSTEM MISSION RELIABILITY

E. E. Lewis
F. Boehm
C. Kirsch
B. P. Kelkhoff

Department of Mechanical Engineering
Northwestern University
Evanston, IL 60208, U. S. A.

ABSTRACT

A Monte Carlo methodology for the reliability simulation of highly redundant systems is presented. Two forms of importance sampling, forced transitions and failure biasing, allow large sets of continuous-time Markov equations to be simulated effectively and the results to be plotted as continuous functions of time. A modification of the sampling technique also allows the simulation of both nonhomogeneous Markov processes and of nonMarkovian processes involving the replacement of worn parts. A number of benchmark problems are examined. For problems with large numbers of components, Monte Carlo is found to result in decreases in computing times by as much as a factor of twenty from the Runge-Kutta Markov solver employed in the NASA code HARP.

1. INTRODUCTION

There is an increasing need to predict mission unreliability and related parameters for systems exhibiting very low rates of failure. Typically, such systems are designed in configurations with many component redundancies and are organized in such a manner that there are component dependencies in the forms of standby subsystems, shared-load components, and shared repair or fault handling faculties. The utility of probabilistic analysis based on combinatorial techniques may be extremely limited. In contrast, such systems may often be modeled as continuous-time Markov processes, particularly if the models are generalizable to include nonhomogeneous Markov processes.

While Markov processes may be an excellent modeling tool, difficulties arise in carrying out computations, particularly in models that are too large or complex to treat my conventional analytical means. As n, the number of components, increase the 2^n explosion of states means that very large systems of coupled differential equations must be solved. Moreover, these equations tend to be very stiff since the time constants involved may range from fault occurrences that are rare events even over weeks or months to fault handling mechanisms that take place in small fractions of a second. As a result, the number of distinct components that can be treated is severely restricted if deterministic methods are employed. If the time constants fall into two widely separated time domains, behavioral decomposition (Bavuso, et al., 1987) may be employed to treat the short time constant events as instantaneous changes of state. But difficulties may then arise when there is inadequate separation in the magnitudes of the time constants.

We have found that Monte Carlo methods may be an effective tool for treating the simulation of systems having highly redundant configurations of components (Lewis and Boehm, 1984; Lewis and Tu, 1986; Boehm, et al., 1988). Regardless of whether component dependencies are present, modeling the system as a continuous-time Markov process allows the average number of event samplings required per trial to be reduced to only slightly more than one. More important, however, is the use of a form of importance sampling that we refer to as forced transitions, to ensure that a substantial fraction of the independent trials will contribute to the tally of the system unreliability. Monte Carlo analysis may be further refined with a second form of importance sampling, referred to as failure biasing, that has the potential for eliminating the approximations inherent in behavioral decomposition. Finally, Monte Carlo tallies may be constructed to yield more than the traditional single answer results; tallies of reliability or other quantities of interest may be generated as continuous functions of time to provide more physical insight into the meaning of the results.

2. MONTE CARLO FORMULATION

For purposes of the Monte Carlo simulation the nonhomogeneous Markov equations are converted to semi-Markov equations. If $p_k(t)$ represents the probability that the system is in state k at time t, then

$$\frac{\partial}{\partial t}p_k(t) = -\gamma_k p_k(t) + \sum_{k'} q(k|k',t)\,\gamma_{k'} p_{k'}(t) ,$$

where the initial conditions are given by

$$p_k(0) = \delta_{k0} .$$

If $\lambda_{jk}(t)$ is the transition rate from state k to state j, then the net transition rate out of state k is

$$\gamma_k = \sum_j \lambda_{jk}(t) ,$$

and the quantity

$$q(k|k',t) = \frac{\lambda_{kk'}(t)}{\gamma_{k'}}$$

is the conditional probability of arrival in state k, given a transition out of state k' at t.

In a Markov process the self-transition rates λ_{kk} vanish. However since effective Monte Carlo sampling requires the values of γ_k appearing in the Markov equation to be independent of time, we treat nonhomogeneous Markov processes by forcing the transition rates γ_k to have positive value that are independent of time. This is accomplished by defining a fictitious self-transition rate

$$\lambda_{kk}(t) = \gamma_k - \sum_{j\neq k} \lambda_{jk}(t) ,$$

where γ_k is taken to be sufficiently large that $\lambda_{kk}(t)$ will remain nonnegative. In cases where the transition rates either remain constant or increase with time this may be achieved by letting

$$\gamma_k = \sum_{j\neq k} \lambda_{jk}(T) ,$$

where T is the mission time.

2.1. Analog Monte Carlo

Analog Monte Carlo trials are performed as indicated in Figure 1. The times to the successive transitions are determined by setting the cumulative distribution function

$$F(t|t',k') = 1 - e^{-\gamma_{k'}(t-t')}$$

equal to a uniformly distributed random number ξ and solving for t,

$$t = t' - \frac{1}{\gamma_k} \ln(1 - \xi) .$$

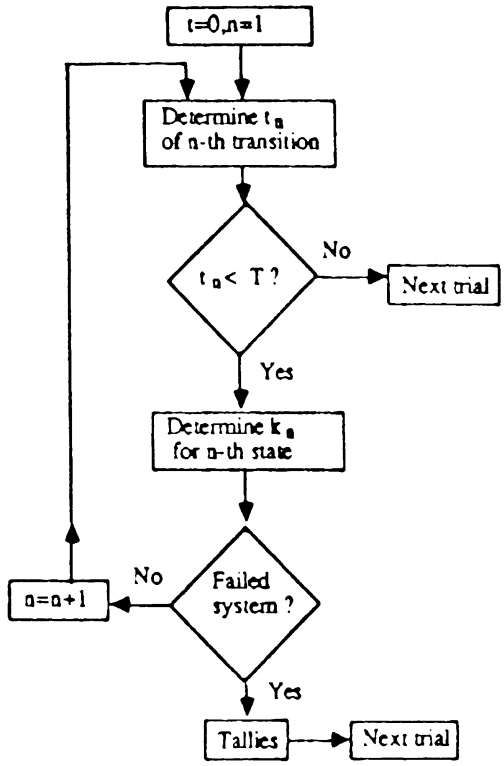

Figure 1: Monte Carlo Trial Procedure for a Design Life T

The new system state is determined by generating a second uniformly distributed random number ζ and choosing the state for which

$$\frac{\lambda_{kk'}(t)}{\gamma_{k'}} \le \zeta \le \frac{\lambda_{k+1,k'}(t)}{\gamma_{k'}} .$$

This procedure is repeated until the mission time is exceeded or the system reaches an absorbing (i.e., failed) state. At any given time the unreliability is just the fraction of trials that have reached failed states.

2.2. Forced Transitions

In highly reliable systems the foregoing algorithm will in most cases require only one sampling per history since the first state transition is not likely to occur until $t > T$. This also means that only a very small fraction of the histories will contribute to the tally, and as a result the variance in the result will tend to be large. To circumvent this difficulty we may modify the distribution of the time to the next transition to force additional transitions within the time interval $0 < t < T$ while modifying the tally such that the results are unbiased. The modified cumulative distribution is

$$\tilde{F}(t|t',k') = \frac{1 - e^{-\gamma_{k'}(t - t')}}{1 - e^{-\gamma_{k'}(T - t')}} \quad , \quad t' < t < T$$

With the uniformly distributed random number ξ , the time of the next transition is then determined from

$$t = t' - \frac{1}{\gamma_{k'}} \ln\left\{1 - \xi\left[1 - e^{-\gamma_{k'}(T - t')}\right]\right\} .$$

To obtain an unbiased result a weight w_i is attached to each trial and initialized at $w_i = 1$. Each time that forced transition sampling is performed the weight is modified by

$$w_i \rightarrow w_i\left[1 - e^{-\gamma_k(T - t')}\right] .$$

The tally for the unreliability is then

$$u_T = \frac{1}{N} \sum_{t_i \le T} w_i ,$$

with a sampling variance given by

$$S^2(u_r) = \frac{1}{N - 1} \sum_{n=1}^{N} [w_n - u_r]^2 .$$

2.3. Failure Biasing

Forced transitions assure that faults will occur in a substantial fraction of the Monte Carlo trials. However in some situations the sampling may remain poor. In mechanical systems, for example, repair rates typically are orders of magnitude larger than component failure rates. Likewise, in avionic systems electronic fault handling systems result in state transition rates that are much faster than the rates at which failures are induced into the system. To further enhance the effectiveness of the Monte Carlo simulation the fraction of smaller probability failure transitions may be increased by the use of a second variance reduction technique which we refer to as failure biasing.

In failure biasing the transition probabilities q(k|k') are modified to increase the ratio of failures to other events such as successful fault handlings. We first divide the transitions out of state k' into to sets; Λ includes those resulting from component failures and R those resulting from successful repair or fault handling. We may then write

$$\gamma_k = \sum_{j \in \Lambda_k} \lambda_{jk}(t) + \sum_{j \notin R_k} \mu_{jk} \ .$$

We require that some fraction x of the transitions come from the set L. The biased transition probabilities are then

$$\tilde{q}(k|k') = \frac{q(k|k')}{\sum_{k'' \in \Lambda} q(k''|k')} x \ , \qquad k \in \Lambda \ ,$$

and

$$\tilde{q}(k|k') = \frac{q(k|k')}{\sum_{k'' \in R} q(k''|k')} (1 - x), \qquad k \in R$$

To maintain unbiased results the trial weight is modified by

$$w_i \rightarrow w_i \frac{1}{x} \sum_{k'' \in \Lambda} q(k''|k')$$

for component failures and

$$w_i \rightarrow w_i \frac{1}{(1 - x)} \sum_{k'' \in R} q(k''|k')$$

for repair. In using failure biasing we typically choose x to be between 0.5 and 0.6; studies of model problems have indicated that values as high as 0.75 may be used before one begins to observe the increases in the sample variance that arise from improbable but very high weight histories (Kirsch, 1988).

3. APPLICATIONS

Two classes of problems are considered in order to examine the accuracy and efficiency of Monte Carlo methods The first consists of simple hybrid redundant systems for which we have also obtained analytical solutions. By varying the ratio of failure to fault handling rate the ability of the variance reduction methods to provide accurate simulations can be determined for systems with very small failure probabilities. In the second class of problems are included two benchmark configurations for which computing times and deterministic solutions have been obtained using the NASA Hybrid Automated Reliability Predictor (HARP) (Bavuso, et al., 1987a, 1987b).

Behavioral composition is employed in the HARP code to separate fault/error-

handling models from the fault occurrence models. The code includes the capability for treating a variety of error handling models, while fault occurrence is modeled as an nonhomogeneous continuous-time Markov process. The imperfect coverage fault/error handling models are reduced to a set of transition probabilities, allowing the entire system to be treated with nonhomogeneous Markov equations in which only the longer time constants of fault occurrence appear. The HARP code solves the Markov equations by the Runge-Kutta method.

3.1. Hybrid Model Problem

We consider a simple hybrid (Lewis and Tu, 1986; Bavuso, et al., 1987) system for which we have obtained analytical solutions elsewhere (Kirsch, 1988). It consists of three units in a majority vote configuration with one spare. Each of the units including the spare has a constant failure rate l, where it is assumed that the spare can not fail until it is switched in. Coverage of the fault by switching in the spare takes place with a constant rate u. The ten hour mission system failure probability is shown in Table 1 over a large range of parameters, with λ and υ given in hrs^{-1}. The ability of Monte Carlo simulation with variance reduction to provide accurate estimates of very small failure probabilities is clearly illustrated.

Table 1: Model Problem Comparison of Analytical and Monte Carlo Unreliability with N = 1000

λ	υ/λ	exact solution	Monte-Carlo solution	COV	relative error
10^{-2}	10^{5}	$0.2464\ 10^{-2}$	$0.2507\ 10^{-2}$	0.08827	0.01745
10^{-3}	10^{6}	$0.2999\ 10^{-5}$	$0.3073\ 10^{-5}$	0.09123	0.02471
10^{-4}	10^{7}	$0.3592\ 10^{-8}$	$0.3667\ 10^{-8}$	0.07626	0.02088
10^{-5}	10^{8}	$0.8997\ 10^{-11}$	$0.9136\ 10^{-11}$	0.02628	0.01545
10^{-5}	10^{9}	$0.6299\ 10^{-13}$	$0.6371\ 10^{-13}$	0.01476	0.01143

3.2. Three-Processor Two-Memory One-Buss System

The detailed problem specification for the 3-processor, 2-memory 1-buss system is given elsewhere (Bavuso, et al., 1987). Briefly, the system is modeled by the Markov diagram shown in Figure 2, where λ, υ and σ represent the processor, memory and buss failure rates. The three numbers associated with each Markov state are the number of operational processors, memory units and buses, respectively, and F1 through F3 are states of system failure. Not shown are the direct transitions from each of the states to system failure that result from near-coincidence and single point failures. The relative frequencies of such failures are determined from the AIRES fault/error handling model (Bavuso, et al., 1987) and appear in the Markov equations as modifications of the state transition probabilities.

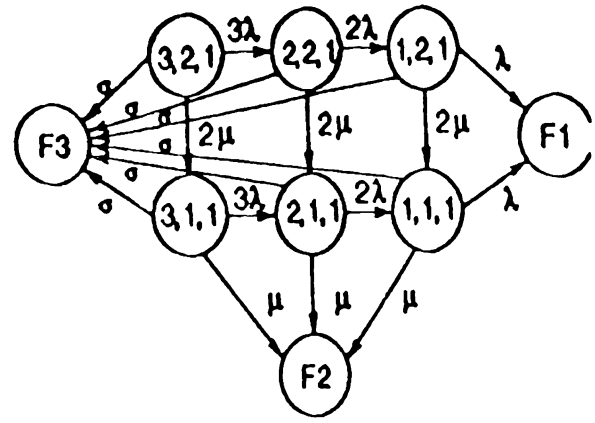

Figure 2: Markov Representation of the 3-Processor, 2-Memory, 1-Buss System

Figure 3 shows Monte Carlo results for the system unreliability over a mission time of ten hours. The data, given in Bavuso, et al., 1987, is for time-independent fault occurrence rates. The three lines correspond to the point estimate and the 68% confidence interval. The Monte Carlo results shown in Fig. 4 are for the same

system, but with Weibull distributions for fault occurrence rates; these have moduli of m = 2.5. In both cases the Monte Carlo simulations consisted of 10,000 trials.

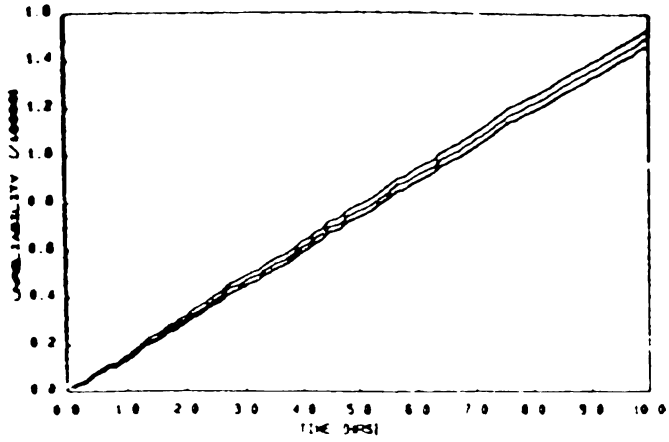

Figure 3: Unreliability vs. Time for the 3-Processor, 2-Memory, 1-Buss System with Constant Failure Rates

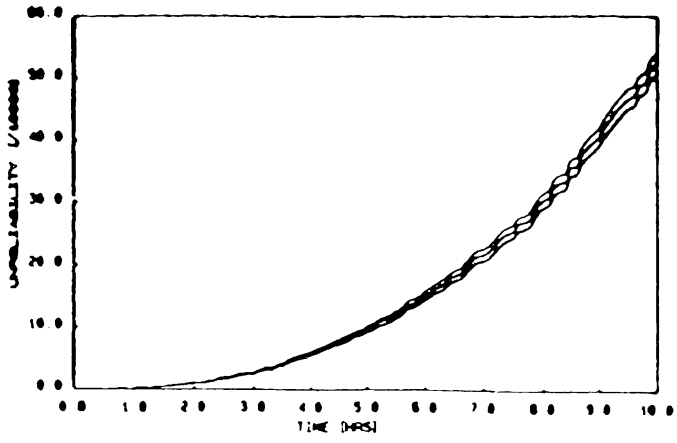

Figure 4: Unreliability vs. Time for the 3-Processor, 2-Memory, 1-Bass System with Incresing Failure Rates

Table 2 indicates that the results from the Monte Carlo and HARP calculations are in excellent agreement; all CPU times are on a VAX 11/785. The Monte Carlo simulations also provide reasonable estimates of the smaller probabilities corresponding to particular failure modes. As an extreme example, the near-coincidence failure probability given as $2.79\ 10^{-11}$ by HARP is estimated as $1.27\ (\pm 1.26)\ 10^{-11}$. Since this few-component problem can be reduced to a set of only six nonabsorbing Markov states, it is not surprising that the Monte Carlo simulations are longer running. It is instructive to note, however, that even for small problems the running times are comparable when Weibull distributions are employed.

Table 2: Ten Hour Mission Unreliability for a 3-Processor 2-Memory 1-Buss System

Constant Failure Rates	Monte Carlo	HARP
Unreliability	$1.498\ (\pm 0.034) 10^{-4}$	$1.521\ 10^{-4}$
CPU sec.	56	~6
Weibull Failures	Monte Carlo	HARP
Unreliability	$4.789(\pm 0.158)\ 10^{-3}$	$4.783\ 10^{-3}$
CPU sec.	582	796

3.3. Jet Engine Control System

The jet engine controller problem, specified in detail elsewhere (Bavuso, et al., 1987), provides a basis for comparing the Monte Carlo and HARP codes for a system with a larger number of components. The CARE II model (Bavuso, et al., 1987) is used for error/fault-handling. The 20 component system has 171 minimum cut sets and is highly redundant as indicated by the fault tree representation shown in Figure 5. The Monte Carlo results for a 10 hour mission are shown in Figure 6. The Monte Carlo unreliability estimate of 1.073 $(\pm 0.087)\ 10^{-5}$ compares well with the HARP result of $1.112\ 10^{-5}$.

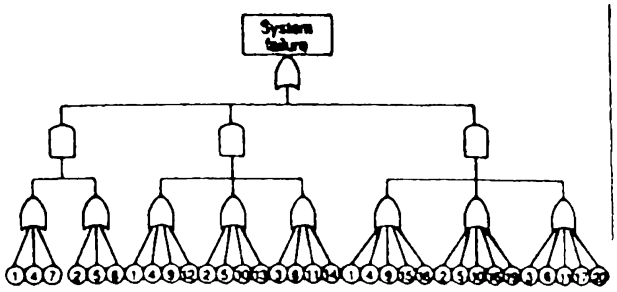

Figure 5: Fault Tree Representation of the Jet Engine Control System

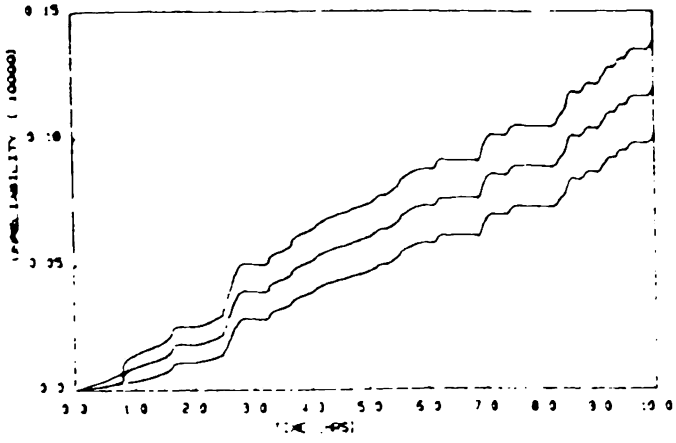

Figure 6: Unreliability vs. Time for the Jet Engine Control System

The time advantages of Monte Carlo simulation become apparent for problems with many Markov states. While the 10,000 history simulation from which the above results were obtained required 20 minutes on the VAX 11/785 the time that would be required by HARP on the same machine is estimated to be of the order of 10 hours. To examine the effect of time-dependent failure rates on the Monte Carlo simulation times the power supply failure rates in the jet engine control were replaced with Weibull distributions with modulus two (Kelkhoff, 1989). This results in less than a 50% increase in the computing time needed to obtain comparable confidence intervals on the unreliability. The Monte Carlo model has also been generalized to allow nonMarkovian as-good-as-new parts replacement on the power supply components. Such modeling increased the computing time by roughly a factor of three over the constant failure rate model but allows problems to be simulated by Monte Carlo that cannot be treated with HARP.

ACKNOWLEDGEMENT

This work was supported in part by Air Force Office of Scientic Research contract 84-0340.

REFERENCES

Bavuso, S. J., Dugan, J. B., Trivedi, K. S., Rothmann, E. M. and Smith, W. E. (1987). Analysis of fault tolerant architectures using HARP. *IEEE Transactions Reliability R-36*, 187-193.

Bavuso, S. J., Dugan, J. B., Trivedi, K. S., Smotherman, B. and Boyd, M. (1987). Applications of the hybrid automated reliability predictor. NASA Technical Paper 2760.

Boehm, F., Hald, U. P., and Lewis, E. E. (1988). Parts renewal in continuous-time Monte Carlo reliability simulation. *Proceedings of 1988 Reliability and Maintainability Symposium*, Los Angeles, California, 345-349.

Kelkhoff, B. P. (1989). Incorporation of preventive maintenance into Monte Carlo reliability analysis. Unpublished project report, Mechanical Engineering Department, Northwestern University, Evanston, Illinois.

Kirsch, C. (1988). Fault/error-handling model incorporation into Markov Monte-Carlo reliability analysis. Unpublished project report, Mechanical Engineering Department, Northwestern University, Evanston, Illinois.

Lewis, E. E. and Boehm, F. (1984). Monte Carlo simulation of Markov unreliability models. *Nuclear Engineering and Design* **77**, 49-62.

Lewis, E. E. and Tu, Z. (1986). Monte Carlo modeling by inhomogeneous Markov processes. *Reliability Engineering* **16**, 277-296.

AUTHORS' BIOGRAPHIES

ELMER E. LEWIS is Professor and Chairman, Department of Mechanical Engineering, Northwestern University. He received the B.S. in engineering physics and the M.S. and Ph.D. in nuclear engineering from the University of Illinois at Urbana. He served as Captain in the U.S. Army and as an Assistant Professor and Ford Foundation Fellow at Massachusetts Institute of Technology before joining Northwestern's faulty.

Elmer E. Lewis
Department of Mechanical Engineering
Northwestern University
Evanston, IL 60208, U.S.A.
(312) 491-3579

FRANZ BOEHM is a staff member at Kernforschungszentrum, Karlsruhe, Federal Republic of Germany. He received the Diplom Ingenieur in nuclear engineering from the University of Stuttgart and the Ph.D. in mechanical engineering from Northwestern University.

Franz Boehm
Kernforschungszentrum, Karlsruhe, IMF IV
Postfach 3640
D-7500 Karlsruhe 1
Federal Republic of Germany

CHRISTOPH KIRSCH is a candidate for the degree Diplom Ingenieur at the University of Stuttgart. In 1987-88 he participated in a student exchange program sponsored by the German Academic Exchange Agency (DAAD) and Northwestern University.

Christoph Kirsch
Hochstetterstr. 19
D-7254 Hemmingen
Federal Republic of Germany

BARBARA P. KELKHOFF is a member of the technical staff at AT&T, Naperville, Illinois. She received the B.S. in mechanical engineering from Bradley University and the M.S. in mechanical engineering from Northwestern University.

Barbara P. Kelkhoff
AT&T Information Systems
1100 E. Warrenville Rd.
Naperville, IL 60566