Reliability Engineering and System Safety 130 (2014) 27–32

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Reliability growth by failure mode removal

Donald P. Gaver*, Patricia A. Jacobs

Operations Research Department, Naval Postgraduate School, Monterey, CA 93943, USA

ARTICLE INFO

Article history:
Received 1 April 2014
Accepted 23 April 2014
Available online 9 May 2014

keywords:
Reliability growth
Infinite server queue
Poisson thinning

ABSTRACT

Modern systems, civilian (e.g. automotive), and military (manned and unmanned aircraft, surface vehicles, submerged vessels), suffer initial design faults or failure modes (FMs), including software bugs, which detrimentally affect the system's reliability and availability. FMs must be removed or mitigated in impact during initial testing, including accelerated testing, in order for the system to meet its reliability requirements and operate satisfactorily in the field. This paper concerns models for reliability growth in which the behaviors of FMs are assumed independent, but of different types. Test effort is guided by prior information, expressed probabilistically, on the random number and tenacities of such FMs that are of various origins in the designs. Estimation of the numbers of FMs that will ultimately activate while in the field is considered here.

Published by Elsevier Ltd.

1. Introduction

Failure mode removal from any system, both hardware and software, is a dynamic uncertain process; see [1–19] for various discussions of problem approaches. In [5,6,8,9,16] an unknown number of FMs are supposed present in the system initially, and the subsequent random times until FM activations are independent and identically distributed; in [8,9] the unknown number of failure modes is assumed deterministic and asymptotic arguments are used for its estimation; in [5,6] the unknown number of FMs has a Poisson distribution, and estimation of the Poisson mean is discussed; [13] includes a Bayesian treatment of the general model; [16] considers a dynamic statistical model for mean number of FMs remaining. The models in [3,4] are widely used and FM activations occur according to a nonhomogeneous Poisson process; a Bayesian treatment of this model appears in Refs. [14,15]. Additional nonhomogenous Poisson process models have been suggested, including that of [17]. In [12] time series are used to summarize software failure data; parameter estimation uses a genetic algorithm; estimation is illustrated with small data sets. In [18,19] a neural network approach is discussed. In [7,9] reliability growth models are suggested for the management of system testing. In [10] a series system of subsystems with resulting FM masking is considered. The goal of failure avoidance, or *system reliability growth* remains a concern to military and civilian system designers, testers and operators; see Toyota automobile accelerator pedal occasional mishaps ([20] and also [21]).

This paper presents an approach to modeling and statistical analysis based on familiar applied stochastic process theory. The model notion is that of identifying failure mode creation and removal with an "infinite server queue", a generalized so-called M/G/∞ system; here M refers to a general Markovian/memoryless "arrival process" of failure modes into a system; G represents the general distribution function of the "residence time", or "service time" in queueing language, of any FM in the system: either until discovered and rectified, or, if not discovered during test, activates in use, thus interrupting field operation usage and possibly causing fatality. The individual FM residence times are here assumed independent and identically distributed; however see [22] for plausible variation. Finally, "∞" refers to the practically infinite, or unlimited number of locations/sites in the system where FMs can reside; cf. [23]; these are "servers" in queueing context as in Ref. [24]. Note that here the items present are all eligible for service/removal when recognized. Later work will recognize congestive servers, and evaluate priority removal. The M-arrival process can include initial numbers of FMs of different types having independent Poisson distributions, with additional FMs that are inadvertently inserted during development according to nonhomogeneous Poisson processes (NHPPs). The assumptions that the unobservable initial number of FMs in a system have Poisson distributions and that the unobservable insertion of additional FMs in a system occur according to NHPPs is convenient and has been made before; see [24–26]. Since the presence of a FM in a system is a "rare event" the assumptions are reasonable, *prima facie*.

The M/G/∞ queue can represent many features encountered in reliability growth data, as has been pointed out by [24–26]. In [24] the NHPP is exploited to describe single-type fault (FM) occurrence; our current results represent realistic recurrence of non-removed

* Corresponding author. Tel.: +1 8316562605; fax: +1 8316562595.
E-mail address: dgaver@nps.edu (D.P. Gaver).

http://dx.doi.org/10.1016/j.ress.2014.04.012
0951-8320/Published by Elsevier Ltd.

FMs. In [25] the discovery of FMs is modeled as occurring according to a NHPP and the times until removal of FMs as independently distributed random variables. The present paper extends the NHPP model of [26] to include fault (FM) type differences plus the realistic probability that actuated FMs are not removed until a removal success. A FM's residence time can include multiple occurrences of a FM due to unsuccessful attempts to remove the FM and may be summarized by a renewal process or even an epidemic process; see [27]. Our present work covers the likely random variation between activation-recurrence rates of different FMs; see (4)–(7) in this paper. These are seen to be natural and realistic extensions of [24–26]. The actual number of FMs is unknown and random, and is realistically controlled by the balance of the arrival and the service or fault-removal process, and so may actually be small, but can grow indefinitely, as in [28]. Note that the mathematical model can be time – or system age – dependent, so different, even new, FM types can be represented during a system design's lifetime.

The model proposed here, and the statistical methods based on it, does not explicitly represent the phenomenon of *mode masking*, meaning that early FM discovery, e.g. of a defect in a vehicle ignition system, or missile launch stage, does not here effect appearance of later FMs that may occur had the early FM not occurred, e.g. in vehicle steering, or missile guidance and detonation. We may view the present model as of one stage, s, of an S-stage series or sequentially operating system. The present model omits desirable mention of Prognostic Health Management (PHM), meaning anticipatory replacement of failure-imminent components or subsystems.

FMs remaining after testing detrimentally affect system field reliability and availability. The purpose of the model and its generalizations is to infer the properties of *FMs remaining* following system testing. The formal model is presented and discussed in Section 2 with examples of behavior that can be represented with the multi-type M/G/∞ queue. Of particular interest are *statistical models* that represent the inherent variation between FMs. A discussion of statistical inference is in Section 3. Section 4 illustrates issues of statistical inference using sample software testing data. The statistical analysis suggests that several different models summarize the data well and that more and extended software testing would be prudent; models with more parameters appear unneeded to summarize these data. The paper ends with conclusions in Section 5.

2. Model for failure mode increase and decrease

Let

- $\lambda_k(t)$ = Random arrival rate of failure modes (FMs) of type k into the system at time t, where $\lambda_k(t)$is the rate of a time-dependent Poisson process. These lie dormant until "flare up" or activation. In other, and subsequent work we can allow FMs to issue warnings or diagnostic symptoms that, if detected, can forestall serious failure. Such FMs can result from human intervention to repair others. $F_k(\tau) = P\{\boldsymbol{T}_k \le \tau\}$ = Probability distribution of $\boldsymbol{T}_k$, the random time until the activation of a single type k FM.

Initially it is assumed that such times can repeat themselves, to represent activations that occurred repeatedly but have not been successfully removed; the times between activations being independent and identically distributed. A special case of inter-activation time distribution is the exponential distribution, $\exp(\mu_k)$,

$$F_k(\tau, \mu_k) = 1 - \exp\{-\mu_k \tau\} \tag{1}$$

Note that the inter-activation time distribution function $F_k(\tau)$, and the arrival rate, $\lambda_k(t)$, can both be affected by environmental influences, including human maintenance or operator, by incorporation of suitable parameter sets and variables. Such important effects are not treated here; they are left for later work.

Next,

- ρ_k = Probability that a FM of type k is removed on any activation. This parameter is initially assumed constant no matter how many responses to activations have occurred; it is a candidate for modification to represent learning.
- $A_k(\tau)$ = Event that a failure mode of type k, is active, i.e. a latent failure, in the system at time τ after it "arrives" in a design or a copy thereof.

Then

$$P\{A_k(\tau)|F_k(\bullet), \rho_k\} = \sum_{n=0}^{\infty} \left[F_k^{*n}(\tau) - F_k^{*(n+1)}(\tau)\right](1-\rho_k)^n \tag{2}$$

where F^{*n} is the n-fold convolution of the distribution F with itself. This simply says that a fault that arrives in the system at $t = 0$ has (independently) activated any number, n, times but has not been removed by time τ. In the special $\exp(\mu_k)$ case

$$P\left\{A_k(\tau)|\exp(\mu_k), \rho_k\right\} = \sum_{n=0}^{\infty} e^{-\mu_k \tau} \frac{(\mu_k \tau)^n}{n!}(1-\rho_k)^n = e^{-\mu_k \rho_k \tau} \tag{3}$$

Following [7], assume μ_k is a realization of independent identically distributed random variables with distribution function $H_k(\mu_k) = P\{\boldsymbol{\mu}_k \le \mu_k\}$; that is, while each FM has independent exponential times between activations, different FMs have different mean inter-activation times drawn from a mixing distribution, $H_k(\bullet)$. From (3), this then implies that

$$E\left[e^{-\boldsymbol{\mu}_k \rho_k \tau}\right] = \int_0^{\infty} e^{-\mu_k \rho_k \tau} dH_k(\mu_k) = \tilde{H}_k(\rho_k \tau), \tag{4}$$

where $\tilde{H}_k(s)$ is the Laplace-Stieltjes transform of the distribution function $H_k(\mu_k)$ evaluated at $s = \rho_k \tau$.

We propose two different forms for the mixing distribution, H_k

(A) Classical gamma.
(B) Positive stable law ([27]).

First, a simple explicit result for the transform of the *gamma distribution* assumed by [29] with scale β_k and shape α_k is

$$\tilde{G}_k(\rho_k \tau) = \left(1 + \frac{\rho_k \tau}{\beta_k}\right)^{-\alpha_k} \tag{5}$$

Next, the positive *stable law* has relevant transform, for shape parameter $0 < \alpha_k < 1$,

$$\tilde{S}_k(\rho_k \tau) = \exp\left\{-\left(\frac{\rho_k \tau}{\beta_k}\right)^{\alpha_k}\right\} \tag{6}$$

notationally (only) matching the stable scale and shape parameters to those of the gamma for $0 < \alpha_k < 1$.

It is evident that (5) and (6) represent the distribution of residence time in the system of corresponding type k fault. Let H_k here represent either G_k or S_k: the probability the residence time is less than or equal to τ is

$$P(A_k(\tau)^c) = 1 - \tilde{H}_k(\rho_k \tau) \tag{7}$$

Put

- $\boldsymbol{N}_k(t)$ = Random number of FMs, that *activate and are removed* from the system during exposure time t, i.e. within $(0, t]$. Note that this includes those initially within the system plus those that are introduced thereafter.
- $\boldsymbol{R}_k(t)$ = Random number of *native* FMs that either have not yet activated or have activated but are not (yet) removed during exposure time t.