Journal of Computer and System Sciences 80 (2014) 323–345

Contents lists available at SciVerse ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

An approach for lifetime reliability analysis using theorem proving

Naeem Abbasi *, Osman Hasan, Sofiène Tahar

Dept. of Electrical & Computer Engineering, Concordia University, 1455 de Maisonneuve West, Montreal, Quebec, H3G 1M8, Canada

ARTICLE INFO

Article history:
Received 1 November 2010
Received in revised form 20 February 2012
Accepted 4 April 2013
Available online 7 June 2013

Keywords:
Reliability analysis
Lifetime analysis
Failure rate
Hazard function
Fractile function
Statistical properties
Random variables
Formal methods
Theorem proving
HOL

ABSTRACT

Recently proposed formal reliability analysis techniques have overcome the inaccuracies of traditional simulation based techniques but can only handle problems involving discrete random variables. In this paper, we extend the capabilities of existing theorem proving based reliability analysis by formalizing several important statistical properties of continuous random variables like the second moment and the variance. We also formalize commonly used concepts about the reliability theory such as survival, hazard, cumulative hazard and fractile functions. With these extensions, it is now possible to formally reason about important measures of reliability (the probabilities of failure, the failure risks and the mean-time-to failure) associated with the life of a system that operates in an uncertain and harsh environment and is usually continuous in nature. We illustrate the modeling and verification process with the help of examples involving the reliability analysis of essential electronic and electrical system components.

1. Introduction

Tragedies such as the industrial accident in the union carbide pesticide plant in Bhopal, India [1], the space shuttle Columbia and Challenger accidents [2], and the high-speed train accident near the village of Eschede in Lower Saxony in Germany [3] all highlight the importance of design reliability in various disciplines of engineering. The reliability of a system is defined as the probability that it will adequately perform its specified purpose for a specified period of time under the specified environmental conditions [4]. The two most popular representations of the distribution of the lifetime of a system are the survival function and the hazard function [4]. The survival function describes the probability that a system is functioning at any time t, and the hazard function describes the failure risk at a time t.

Traditionally, reliability analysis has been done using either paper-and-pencil or simulation based approaches. In engineering applications, the paper-and-pencil approach very quickly becomes impractical because of the amount of detail involved. Simulation based reliability analysis is popular because of the availability of a number of automated tools. Unfortunately, the simulation based analysis is neither accurate nor can it truly model random behavior. Computer simulations rely on floating-point numbers based representation of system parameters which can lead to errors in reliability analysis and thus can have costly consequences. For example, the floating point bug in Intel Pentium 5 was related to a few missing entries in the lookup table used by the digital divide operation algorithm. This resulted in rare round-off errors. Intel eventually had to recall the flawed chips at a cost of over 475 million dollars [5]. Moreover, simulation based techniques use

* Corresponding author.
E-mail addresses: n_ab@ece.concordia.ca (N. Abbasi), o_hasan@ece.concordia.ca (O. Hasan), tahar@ece.concordia.ca (S. Tahar).

0022-0000/$ – see front matter
http://dx.doi.org/10.1016/j.jcss.2013.05.002

pseudo random number generators for simulating the random behavior and require a large number of computing resources. Another problem with the simulation based analysis is the time needed to compile and interpret simulation results. Some simulation based methods [6] utilize computation tools, such as MATLAB [7], which supports arbitrarily large base sizes for number representation and helps in controlling computational errors. This, however, comes at the cost of significant increase in simulation times. For some applications, such as communication networks, the simulation softwares often make overly simplified assumptions which are not realistic [8] or use different levels of details [9] and lead to results that do not match real world measurements. Formal method-based techniques are 100% accurate and allow the modeling and analysis of true random behavior and thus provide an alternative approach for reliability analysis of the critical parts of a system.

Formal techniques that analyze system reliability using probabilistic models, such as probabilistic model checking, are accurate; however, they cannot effectively handle properties that summarize the statistical behavior of a lifetime distribution, such as its moments and variance. Moreover, similar to traditional model checking techniques, probabilistic model checking techniques also suffer from the state space explosion problem, and, therefore, only a small set of reliability analysis problems can be handled using these techniques. With techniques based on theorem proving, it is possible to accurately deal with complexity and reason about reliability analysis related probabilistic and statistical properties for large sized engineering problems. These techniques, however, are very often interactive and require a formalized infrastructure for reasoning.

The state of the art in theorem proving based probabilistic analysis consists of a formalization of measure and probability theories [10], discrete and continuous random variables and their probabilistic [11] and expectation properties [11,12]. Some of these foundations have been utilized to asses reliability aspects involving discrete random variables [13] and expectation properties of continuous random variables [12]. Despite these efforts, there are many reliability aspects that cannot be reasoned about in a mechanical theorem prover, namely, the higher-order moments, variance and the concepts of survival function, hazard function, cumulative hazard function and fractile function. In this paper, we formalize these reliability fundamentals by building upon the existing probability [10–12] and reliability [14] theory foundations. Our first contribution is to formally verify a general expression that facilitates reasoning about the second moment of bounded continuous random variables that ranges over the interval $[a, b]$.

$$E\left[X^2\right] = \lim_{n\to\infty}\left[\sum_{i=0}^{2^n-1}\left(a+\frac{i}{2^n}(b-a)\right)^2 \mathbb{P}\left\{a+\frac{i}{2^n}(b-a) \leqslant X < a+\frac{i+1}{2^n}(b-a)\right\}\right] \tag{1}$$

where E is the expectation operator, X is the bounded random variable, and $\mathbb{P}$ is the probability measure. $E[X^2]$ represents the expectation of X squared with respect to $\mathbb{P}$. The initial verification effort involved in proving the correctness of this general result in an interactive theorem proving environment is quite significant. However, it makes the interactive verification of statistical properties of specific random distributions less tedious to handle. This is mainly because the reasoning involved in the verification of statistical properties now involves the concepts of summation of sequences and their limits rather than relatively more involved concepts from set, measure and Lebesgue integration theories.

The second main contribution of this paper is that we utilize the general expression of Eq. (1), along with a similar expression for unbounded random variables, given in [14], to verify several important statistical properties of random variables that are commonly used in reliability analysis. For example, the second moment and variance relations for the Uniform, Triangular, and Exponential random variables are verified. When short term lifetime behavior of a system is of interest uniform distribution is preferred. When very little information about the lifetime behavior of a system is known (such as the rough estimates of the minimum and maximum life of a system, as is sometimes the case in the initial planning and design stages), reliability engineers prefer triangular random distribution. And finally, exponential distribution accurately models constant failure rate behavior, which is the most commonly used distribution for lifetime modeling of electronic and electrical components of a system.

Different lifetime distribution representations have been used in the past depending upon the specific needs of the problem of analyzing lifetime. For example, sometimes the probability of failure is of interest at a certain time (survival function), whereas, in another application such as in planning for serviceability and maintainability of a system, the total amount of risk associated with a system up to a given time (cumulative hazard function) may be required [15]. Two commonly used important reliability properties of survival function and hazard function have already been formalized in [14]. We add to these two more equally important lifetime distribution representations of cumulative hazard function and the fractile function. Cumulative hazard function gives the amount of risk associated with a system up to a given time while the fractile functions allow reasoning about times for a given probability of failure [4]. The survival function $S_X(x)$ is defined as:

$$S_X(t) = 1 - F_X(t) \tag{2}$$

where $F_X(x)$ is the cumulative distribution function of the random variable X. The hazard function, $h_X(t)$, is defined as:

$$h_X(t) = -\frac{\frac{dS_X(t)}{dt}}{S_X(t)} = \lim_{h\to 0}\frac{S_X(t) - S_X(t+h)}{hS_X(t)} \tag{3}$$