Reliability Engineering and System Safety 150 (2016) 89–95

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Replacement policies for a parallel system with shortage and excess costs

Xufeng Zhao [a], Mingchih Chen [b,*], Toshio Nakagawa [c]

[a] Department of Mechanical and Industrial Engineering, Qatar University, Doha 2713, Qatar
[b] Graduate Institute of Business Administration, Fu Jen Catholic University, New Taipei City 24205, Taiwan
[c] Department of Business Administration, Aichi Institute of Technology, Toyota 470-0392, Japan

ARTICLE INFO

Article history:
Received 25 June 2015
Received in revised form
29 November 2015
Accepted 20 January 2016
Available online 4 February 2016

Keywords:
Parallel system
Age replacement
Random replacement
Working times
Parallel computing

ABSTRACT

It has been assumed in most maintenance models that (i) the number of units for a parallel system can be predefined precisely, (ii) maintenance cost after failure should be avoided, and (iii) age replacement can always be performed at its optimized times. However, these assumptions are challenged from practical perspectives in this paper. For this purpose, shortage and excess costs, which claim that replacement done too early before failure involves a waste of operation, are introduced into replacement models, such as replacement plans with time T or distribution $G(t)$ for constant n or random N of units. Furthermore, the number N of working times and the number n of available units for a scheduled replacement time T are discussed. For each model, the expected replacement cost for one cycle and expected replacement cost rate for a long run are optimized, and numerical examples are given.

1. Introduction

High system reliability can be achieved through redundancy. The most typical model is a standard parallel system which consists of n identical units. The redundant system was originally shown to have the capacity to operate for an expected mean time by either increasing the number of units or changing the replacement time [1]. The reliabilities of many redundant systems have been computed and summarized [2]. A variety of redundant systems with multiple failure modes and their optimization problems were discussed [3]. The reliabilities of parallel systems with dependent failures among components [4] and the optimization methods of redundancy allocation for series–parallel systems were studied [5].

However, little research has been done on replacement policies for a parallel system by considering the following newly proposed problems:

(i) For a large-scale parallel system, it is difficult or perhaps impossible to predefine the exact number of units needed to operate jobs. When we make decision to replace this large system preventively, its number of units needed is the random variable that should be estimated rather than a constant value given in advance. Such a stochastic phenomenon has been mentioned for the order statistics in theory, assuming the sample size of components for parallel and series systems to be random variables [6]. Recently, an asymptotic method of computing the mean time to system failure [7] and optimization of replacement times for a random-sized parallel system were obtained [8]. In [8], it was also indicated that the model could be applied for maintaining an aircraft fuselage with multiple site cracks around rivets, where the number of rivets is very large.

(ii) It has been established for age replacement in reliability theory that the operating unit should be replaced only at failure when the random failure has an exponential distribution [1,9]. For example, suppose in (3.9) [9] that $F(t)=1-e^{-\lambda t}$, then the left-hand side of (3.9) equals to 0, which is less than the cost ratio on the right side, that is, optimal replacement time T^* becomes infinite. For an asymptotic model (11) in [10], a finite replacement time T_x also does not exist when the failure time is exponentially distributed. To model the above policy in which a finite replacement time exists for an exponential failure time, two types of costs based on the scales of replacement times, i.e., shortage cost incurred from failure for the carelessly scheduled replacement and excess cost caused by waste of replacement performed too early before failure, have been proposed [11,12]. Applying these two costs to parallel and standby systems, the optimal number of units that should be available to operate a single work or several tandem works were discussed [13,14].

(iii) It has been assumed in most models that systems are inspected and then repaired or replaced at some deterministic times, e.g., planned time t^*, periodic times n^*t, or a condition-based level

* Corresponding author.
E-mail address: 081438@mail.fju.edu.tw (M. Chen).

http://dx.doi.org/10.1016/j.ress.2016.01.008
0951-8320/

δ^*, however, any maintenance performance would be limited by the factor of work cycles operated by the unit. For example, when the system is running, it is impossible or impractical to suspend job operations for inspection or replacement actions due to production losses [15,16]. When a unit is replaced only at random times, the properties of replacement policies between two successive failures and replacement times with arbitrary distributions have been investigated [17].

In this study, the above three aspects are modeled simultaneously into preventive replacement policies for an operating parallel system. A potential application is to make checkpoint schedules for a parallel computing in which a large number of calculations are carried out simultaneously in information systems [18,19], which will be addressed in concluding remarks. Therefore, based on the outlined assumptions (i)–(iii), the main purpose of this paper is to model replacement policies for a parallel system with shortage and excess costs, considering replacement plans with time T or distribution $G(t)$ for constant n or random N of units.

The remainder of this paper is organized as follows: Section 2 gives an age replacement policy that is performed at planned time T for a parallel system, where the number of units is predefined as constant values n and estimated as the random variable N with a truncated Poisson distribution. In Section 3, we suppose that the replacement time T is scheduled randomly, which follows a general distribution $G(t)$ with finite mean $1/\theta$. Optimal T and mean $1/\theta$ of $G(t)$ for replacement times are discussed analytically in Sections 2 and 3. Section 4 describes two other models: optimization of working times N when the parallel system operates for jobs with successive times, and optimal number n of units that should be available when constant T is given. Finally, conclusions of the paper and a potential application of the model are provided in Section 5.

2. Replacement with time T

2.1. Constant n of units

Consider a parallel system with number n of units, in which n is predefined as constant values such that $n = 1, 2, \ldots$. Each unit has an independent and identical failure distribution $F(t)$ with a density function $f(t)$ and finite mean $1/\lambda$ $(0 < \lambda < \infty)$, and the system fails when all units have failed [11]. If the failure time of the system is assumed to be a random variable X, then the probability that it is still operating at time t is

$$\Pr\{X > t\} = 1 - F(t)^n. \tag{1}$$

It is supposed that the system is replaced preventively at planned time T $(0 < T \le \infty)$ or correctively at failure X, whichever occurs first. Then, the mean time to replacement is

$$\mu_T = T[1-F(T)^n] + \int_0^T t\, \mathrm{d}F(t)^n = \int_0^T [1-F(t)^n]\, \mathrm{d}t. \tag{2}$$

We introduce the following two types of costs that are the functions of times for preventive and corrective replacement policies:

(i) If replacement time T is planned too early prior to failure time X, a waste of operation cost $c_E(X-T)$ is incurred because the system might run for an additional period of time to complete critical operations.

(ii) If replacement time T is planned too late after failure time X, a great failure cost $c_S(T-X)$ is incurred due to the delay in time of the carelessly scheduled replacement. We call these two costs *excess cost* $c_E(t)$ and *shortage cost* $c_S(t)$ in the following discussions. Obviously, when $c_E(t) \equiv c_E$ and $c_S(t) \equiv c_S$, c_E and c_S become the costs for preventive replacement before failure and corrective replacement after failure, where c_E should be less than c_S. Next, how replacement time T^* can be determined in advance becomes the problem that remains.

Under these conditions, the expected replacement cost for one cycle is

$$C_1(T;n) = \int_T^\infty c_E(t-T)\, \mathrm{d}F(t)^n + \int_0^T c_S(T-t)\, \mathrm{d}F(t)^n. \tag{3}$$

Thus, because the mean time to replacement is given in (2), the expected replacement cost rate for a long run is

$$C_2(T;n) = \frac{\int_T^\infty c_E(t-T)\, \mathrm{d}F(t)^n + \int_0^T c_S(T-t)\, \mathrm{d}F(t)^n}{\int_0^T [1-F(t)^n]\, \mathrm{d}t}. \tag{4}$$

Note that when $c_E(t) \equiv c_E$, $c_S(t) \equiv c_S$, and $c_E < c_S$,

$$C_2(T;n) = \frac{c_E + (c_S - c_E)F(T)^n}{\int_0^T [1-F(t)^n]\, \mathrm{d}t}, \tag{5}$$

which agrees with the age replacement model for a parallel system [11]. When $n > 1$, optimal T^* satisfies

$$H(T)\int_0^T [1-F(t)^n]\, \mathrm{d}t - F(T)^n = \frac{c_E}{c_S - c_E}, \tag{6}$$

where

$$H(T) \equiv \frac{nf(t)F(t)^{n-1}}{1-F(t)^n},$$

which increases strictly with T to ∞ [11].

When $n=1$, the model becomes age replacement planned at time T [1,9], in which case, the left-hand side of (6) equals to 0 for $F(t) = 1 - \mathrm{e}^{-\lambda t}$, and optimal policy in (5) is $T^* = \infty$, i.e., replacement should be done only at failure.

We next find analytically optimal T_i^* $(i = 1, 2)$ to minimize $C_i(T;n)$ for given n $(n \ge 1)$ when $c_E(t) = c_E t$ and $c_S(t) = c_S t$. In this case, (3) becomes

$$C_1(T;n) = c_E \int_T^\infty [1-F(t)^n]\, \mathrm{d}t + c_S \int_0^T F(t)^n\, \mathrm{d}t. \tag{7}$$

Clearly,

$$C_1(0;n) \equiv \lim_{T\to 0} C_1(T;n) = c_E \int_0^\infty [1-F(t)^n]\, \mathrm{d}t,$$

$$C_1(\infty;n) \equiv \lim_{T\to\infty} C_1(T;n) = \infty,$$

where $\int_0^\infty [1-F(t)^n]\, \mathrm{d}t$ represents the mean time to failure for the system. Differentiating $C_1(T;n)$ with respect to T and setting it equal to zero,

$$F(T)^n = \frac{c_E}{c_E + c_S}, \tag{8}$$

whose left-hand side increases with T from 0 to 1. Therefore, there exists an optimal T_1^* $(0 \le T_1^* < \infty)$ which satisfies (8).

From (4), the expected cost rate is

$$C_2(T;n) = \frac{c_E \int_T^\infty [1-F(t)^n]\, \mathrm{d}t + c_S \int_0^T F(t)^n\, \mathrm{d}t}{\int_0^T [1-F(t)^n]\, \mathrm{d}t}. \tag{9}$$

Clearly,

$$\lim_{T\to 0} C_2(T;n) = \lim_{T\to\infty} C_2(T;n) = \infty.$$

Differentiating $C_2(T;n)$ with respect to T and setting it equal to zero,

$$\frac{\int_0^T [1-F(t)^n]\, \mathrm{d}t}{1-F(T)^n} - T = \frac{c_E}{c_S}\int_0^\infty [1-F(t)^n]\, \mathrm{d}t, \tag{10}$$

whose left-hand side increases strictly from 0 to ∞. Therefore, there exists a finite and unique T_2^* $(0 \le T_2^* < \infty)$ which satisfies