Reliability Engineering and System Safety 94 (2009) 254–260

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Optimal load distribution in series–parallel systems

Gregory Levitin [a,*], Suprasad V. Amari [b]

[a] Israel Electric Corporation Ltd., P.O. Box 10, Haifa 31000, Israel
[b] Relex Software Corporation, 540 Pellis Road, Greensburg, PA 15601, USA

ARTICLE INFO

Article history:
Received 25 February 2007
Received in revised form
2 March 2008
Accepted 5 March 2008
Available online 18 March 2008

Keywords:
Optimal loading
Multi-state system
Failure rate
Universal generating function
Genetic algorithm

ABSTRACT

This paper presents an algorithm for determining an optimal loading of elements in series–parallel systems. The optimal loading is aimed at achieving the greatest possible expected system performance subject to repair resource constraint. The model takes into account the dependence of elements' failure rates on their load. The optimization algorithm uses a universal generating function technique for evaluating the expected system performance, and a genetic algorithm for determining the optimal load distribution. An illustrative example of load distribution optimization is presented.

1. Introduction

A majority of engineering systems are designed to support varying amounts of load. Examples include cargo trains, trucks, conveyors, electric lines, computer processors, and telecommunication channels. In some cases, the load can also be measured in terms of usage frequency or busy period. The load-carrying levels of a system or an element can be discrete or continuous [1–3]. Examples of discrete load-carrying systems include power-generating units in a thermal power plant and discrete unit industrial conveyors. Examples of continuous load-carrying systems include coal conveyors, cargo trucks and trains, and variable feed cutting tools. For practical purposes, we can consider all of them as discrete load capacity systems. For example, it is appropriate to measure the load on a coal conveyor in terms of integer values of tons per hour.

Many empirical studies of mechanical systems [4] and computer systems [5] have proved that the workload strongly affects the component failure rate. Therefore, it is important to consider the load versus failure rate relationship while performing system reliability evaluation and optimization. We describe these models in Section 2.

In reliability engineering, we encounter two classes of problems that consider the effects of load on component failures.

- *Static load distribution problems*: In static models, the load on a given component is almost constant and does not vary with time or other events. However, we have the control on determining the optimal load on that component. An example of this kind of problem is determining the optimal load on a truck or optimal feed of a cutting tool [2,3].
- *Dynamic load distribution problems*: In these models, the load on a given component may vary with time or other events such as failure of other components. An example of this kind of problem is a well-known load-sharing system where the total load on a system is shared by remaining working components. Applications of load-sharing systems include electric generators sharing an electrical load in a power plant, CPUs in a multi-processor computer system, cables in a suspension bridge, and valves or pumps in a hydraulic system [6].

Both the static and dynamic load distribution problems are important. Although the dynamic load distribution approach has a wide range of applications, the optimal dynamic load distribution problem requires capturing the information on the system state and computing the optimal parameters online. However, the static load distribution problem requires less information and does not require tracking the system state and other events. Hence, the static load distribution problems are more practical and require less record-keeping. In this paper, we considered the static load distribution problem.

1.1. Problem description

We consider a series–parallel system with non-identical elements. Each element is capable of supporting discrete loads. The failure rate of an element depends on the load it supports.

* Corresponding author.
E-mail addresses: levitin@iec.co.il (G. Levitin), suprasad.amari@relex.com (S.V. Amari).

0951-8320/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2008.03.001

Nomenclature

RBD	reliability block diagram
MSS	multi-state system
u-function	universal generating function
pmf	probability mass function
Pr{e}	probability of event e
E	expected system performance
N	number of system elements
T	required repair resource
T^*	available repair resource
G_j	random performance of system element j
$\mathbf{g}_j$	set of possible realizations of G_j
g_{jh}	hth realization of G_j
p_{jh}	$\Pr\{G_j = g_{jh}\}$
L_j	load (performance) of two-state element j in its working state
$\mathbf{L}$	vector of elements' loads $\mathbf{L} = \{L_1,\ldots,L_N\}$
μ_j	repair rate of element j
λ_j	failure rate (hazard rate) of element j
A_j	availability of element j
V	random system performance
v_i	ith realization of V
q_i	$\Pr\{V = v_i\}$
ϕ	system structure function: $V = \phi(G_1,\ldots,G_n)$
$u_j(z)$	u-function representing pmf of G_j
$U(z)$	u-function representing pmf of V
$\otimes$	composition operator over u-functions
$\overset{\varphi}{\phi}(G_i,G_j)$	function representing performance of pair of elements
w	penalty function
ω	penalty value

In general, the failure rate increases with the load. Hence, the increase of load on an element increases the number of failures and the associated repair cost (in the long run, the number of failures per unit time is equivalent to the number of repairs).

On the one hand, the increase in load increases the element's performance in working states. On the other hand, the increase in load increases the overall downtime of the element, which reduces its average performance over a long time. Therefore, the expected element performance can be non-monotonic function of its load. One may want to find the optimal load on each element of the system that provides the maximal expected performance.

As long as all elements in the system are independent, the optimal load of each element can be found independently of other elements in the system. Hence, the problem can be simplified to the optimal load distribution of a single element. However, it is true only if there are no system-level constraints associated with the optimization problem, which is hardly the case. In most practical cases, we may have several system-level constraints such as:

- Limited repair resources.
- Allowed total number of failures (or total repair time).
- Upper and lower limits on the entire system performance.
- Allowed system unavailability.

Depending on the context, the term "*repair resources*" is used for different things. In some literature, it is used to describe the available number of repair personnel. It can also be used for the equipment needed for repairing. Similarly, it can be used for the budget allocated for repairing or the allowed average cost of repairs per unit time. In general, the cost of a repair is proportional to the repair time. If cost of repair per unit is the same for all components, then the "average repair cost per unit time" constraint is equivalent to the constraint on the "average value of total repair time of all components per unit time". For simplicity, we refer to this as total repair time constraint. The proposed method can handle all these constraints. However, in order to simplify the problem description, we considered only allowed total repair time constraint.

In this paper all components are considered to be repairable and the repair of a component starts immediately after its failure. This inherently assumes that there are sufficient repair personnel. This assumption is valid for most systems because the cost of downtime and cost of repair are very high as compared to the cost of repair personnel. Hence, in most cases, reasonably sufficient repairmen are available. If waiting time related to unlimited repairmen is really significant, then it is easy to hire additional repair personnel depending on the number of failed components in the system.

When load on each component is fixed, it is assumed that the failure times in each component occur as per exponential distribution with constant rate. We assume that the failure and repair times of a component are independent and repairs are perfect and do not change the failure pattern of the components. As a result, the failure and repair process of each component follows an alternative renewal process. Within each failure–repair cycle, the failure time follows exponential distribution with a constant rate parameter; we represent the failure distribution using its failure rate (hazard rate). The term "failure rate" used in this paper is not equivalent to the failure intensity function. Strictly speaking, it is numerically equivalent to the conditional failure intensity. It should also be noted that the alternative renewal process is a special type of semi-Markov process; therefore, it is appropriate to represent the state transition times in terms of failure and repair distributions defined over local times (time spent in each state) as opposed to the global times used in the non-homogeneous Poisson processes.

To illustrate the suggested methodology, consider a power plant multi-stage coal feeding system consisting of several conveyors with different load-carrying capacities and failure rates. Any given loading of each conveyor determines its capacity in working state and its failure rate. Having these parameters and the system structure, one can determine the entire system's expected carrying capacity and the expected total repair time (or number of repairs) of the conveyors using the algorithm presented below. Since the numerical algorithm obtains these two indices for any arbitrary conveyor's loading, it may be used in an optimization procedure that seeks for the conveyor's load distribution that maximizes the coal feeding system's expected carrying capacity subject to constraints on the expected total repair time (or cost). The numerical example that considers the feeding system with nine conveyors is presented in Section 6.

2. Load–failure rate relationship models

In order to analyze the optimal load distribution systems, we should consider the relationship between the load and the failure behavior of a component. The accelerated life testing models play an important role in determining the relationships between load and failure rate. There is a huge amount of literature on the