The background of the book cover features a complex, abstract geometric pattern. It consists of numerous small, semi-transparent spheres of varying sizes scattered across the surface. Overlaid on these are several large, semi-transparent wireframe-like structures that resemble three-dimensional polyhedra, such as octahedrons and cubes. These structures are composed of thin grey lines connecting vertices. Some of these vertices are also small spheres, creating a sense of depth and connectivity.

EDITED BY

MANGEY RAM

J. PAULO DAVIM

MATHEMATICS

APPLIED TO ENGINEERING



Mathematics Applied to Engineering

Mathematics Applied to Engineering

Edited by

Mangey Ram

Graphic Era University, Dehradun, India

J. Paulo Davim

University of Aveiro, Portugal



ACADEMIC PRESS

An imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, United Kingdom
525 B Street, Suite 1800, San Diego, CA 92101-4495, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom
Copyright © 2017 Mangey Ram and J. Paulo Davim. Published by Elsevier Ltd.
All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-12-810998-4

For information on all Academic Press publications visit our website at <https://www.elsevier.com/books-and-journals>



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Glyn Jones

Acquisition Editor: Glyn Jones

Editorial Project Manager: Edward Payne

Production Project Manager: Poulose Joseph

Cover Designer: Greg Harris

Typeset by SPi Global, India

List of contributors

- A. Barros** Norwegian University of Science and Technology, Trondheim, Norway
- I. Bolvashenkov** Technical University of Munich, Munich, Germany
- L. Cui** School of Management and Economics, Beijing Institute of Technology, Beijing, China
- P. Do** University of Lorraine, Vandoeuvre-les-Nancy, France
- T. Dohi** Hiroshima University, Higashi-Hiroshima, Japan
- I. Frenkel** Sami Shamoon College of Engineering, Beer-Sheva, Israel
- N. Goyal** Graphic Era University, Dehradun, Uttarakhand, India
- P. Gulia** Indian Institute of Technology Mandi, Kamand, India
- A. Gupta** Indian Institute of Technology Mandi, Kamand, India
- H.-G. Herzog** Technical University of Munich, Munich, Germany
- L. KhvatSkin** Industrial Engineering and Management Department, Beer-Sheva, Israel
- A. Kumar** G. B. Pant University of Agriculture and Technology, Pantnagar, India
- A. Kumar** Graphic Era University, Dehradun, India
- A. Kumar** University of Petroleum & Energy Studies, Dehradun, India
- S. Limon** North Dakota State University, Fargo, ND, United States
- C. Lin** School of Management and Economics, Beijing Institute of Technology, Beijing, China
- H. Okamura** Hiroshima University, Higashi-Hiroshima, Japan
- S. Pant** University of Petroleum & Energy Studies, Dehradun, India
- P.P. Patil** Graphic Era University, Dehradun, India

M. Ram Graphic Era University, Dehradun, Uttarakhand, India

S.B. Singh G. B. Pant University of Agriculture and Technology, Pantnagar, India

O.P. Yadav North Dakota State University, Fargo, ND, United States

H. Yang School of Management and Economics, Beijing Institute of Technology, Beijing, China

About the editors

Mangey Ram received B.Sc. degree in science from Chaudhary Charan Singh University, Meerut, India, in 2000, M.Sc. degree in mathematics from Hemwati Nandan Bahuguna Garhwal University, Srinagar (Garhwal), India, in 2004, and Ph.D. degree major in mathematics and minor in computer science from G. B. Pant University of Agriculture and Technology, Pantnagar, India, in 2008. He has been a faculty member for around 9 years and has taught several core courses in pure and applied mathematics at undergraduate, postgraduate, and doctorate levels. He is currently professor with the Department of Mathematics, Graphic Era University, Dehradun, India. Before joining the Graphic Era University, he was a deputy manager (probationary officer) with Syndicate Bank for a short period. He is editor-in-chief of *International Journal of Mathematical, Engineering and Management Sciences*; executive editor, *Journal of Graphic Era University*, executive associate editor of the *Journal of Reliability and Statistical Studies*, and the guest editor and member of the editorial board of many journals. He is a regular reviewer for international journals, including IEEE, Elsevier, Springer, Emerald, John Wiley, Taylor & Francis, and many other publishers. He has published 100 plus research publications in IEEE, Springer, Elsevier, Emerald, World Scientific, and many other national and international journals of repute and also presented his works at national and international conferences. His fields of research are reliability theory and applied mathematics. He is a member of the IEEE, Operational Research Society of India, Society for Reliability Engineering, Quality and Operations Management in India, International Association of Engineers in Hong Kong, and Emerald Literati Network in the United Kingdom. He has been a member of the organizing committee of a number of international and national conferences, seminars, and workshops. He has been conferred with “*Young Scientist Award*” by the Uttarakhand State Council for Science and Technology, Dehradun, in 2009. He has been awarded the “*Best Faculty Award*” in 2011 and recently “*Research Excellence Award*” in 2015 for his significant contribution in academics and research at Graphic Era University.

J. Paulo Davim received his Ph.D. in Mechanical Engineering in 1997, M.Sc. degree in Mechanical Engineering (materials and manufacturing processes) in 1991, Licentiate degree (5 years) in Mechanical Engineering in 1986, from the University of Porto (FEUP), the Aggregate title (Full Habilitation) from the University of Coimbra in 2005, and a D.Sc. from London Metropolitan University in 2013. He is Eur Ing by FEANI and senior chartered engineer by the Portuguese Institution of Engineers with an MBA and Specialist title in Engineering and Industrial Management. Currently, he is professor at the Department of Mechanical Engineering of the University of

Aveiro, Portugal. He has more than 30 years of teaching and research experience in Manufacturing, Materials and Mechanical Engineering with special emphasis in Machining & Tribology. Recently, he has also interest in Management/Industrial Engineering and Higher Education for Sustainability/Engineering Education. He has received several scientific awards. He has worked as evaluator of projects for international research agencies as well as examiner of Ph.D. thesis for many universities. He is the editor-in-chief of several international journals, guest editor of journals, books editor, book series editor, and scientific advisory for many international journals and conferences. Presently, he is an editorial board member of 30 international journals and acts as reviewer for more than 80 prestigious Web of Science journals. In addition, he has also published as editor (and co-editor) more than 80 books and as author (and co-author) more than 10 books, 60 book chapters, and 400 articles in journals and conferences (more than 200 articles in journals indexed in Web of Science/h-index38+ and SCOPUS/h-index 46+).

Preface

Mathematics is the strength of engineering sciences, and applied mathematics forms the common foundation of all novel discipline as engineering evolves and develops. Everything would be possible with the addition of mathematics to the engineering. The book *Mathematics Applied in Engineering* gripped on a comprehensive range of mathematics applied in various fields of engineering different chores such as acoustics, system engineering, optimization, mechanical engineering, and reliability engineering. Through this book entitled *Mathematics Applied in Engineering* the engineers have to gain a greater knowledge and help them in the applications of mathematics in engineering courses. The book is meant for those who take engineering as a subject of study. The material is intended for an audience at the level of postgraduate or senior undergraduate students.

Mangey Ram

Graphic Era University, Dehradun, India

J. Paulo Davim

University of Aveiro, Portugal

Acknowledgment

The editors acknowledge Elsevier and the editorial team for their adequate and professional support during the preparation of this book. Also, we would like to acknowledge all the chapter authors and the reviewers for their availability for work on this book project.

Mangey Ram
Graphic Era University, Dehradun, India

J. Paulo Davim
University of Aveiro, Portugal

Analysis for a qualification test procedure with FMCIA (finite Markov chain imbedding approach)

1

L. Cui, C. Lin, H. Yang

School of Management and Economics, Beijing Institute of Technology, Beijing, China

1 Introduction

Test procedure is a key step to ensure the product is in good quality before it is provided to the customers. A rigorous test plan can guarantee the quality of the product. But it may enlarge the scale of the test procedure, e.g., involving more samples to be tested, which is cost consumed. On the contrary, although an easy test plan leads the product incline to pass the test procedure, the quality issue may evolve during the post-sale service. Therefore, it is always essential to put up with a proper test procedure. In terms of proper test plan, it means the product can pass the test procedure and is accepted with higher probability but fewer samples to be tested. Thus, in this chapter, we consider to improve the products whenever a failure of test occurs. We put up with some index such as probability of acceptance, expected number of tested samples, and the expected number of improvement actions to evaluate the test procedure. We conduct an analysis for the quality test procedure by employing the FMCIA (finite Markov chain imbedding approach).

The organization of this chapter is as follows: We first conduct a literature review for the FMCIA and illustrate the proposed quality test procedure. After that, we present the mathematical description of the quality test plan proposed in this chapter. Then equations for calculating the indexes of the test procedure, e.g., the probability for passing the test, the expected number of the tests, and the expected number of the improvement actions, are derived. Thereafter, we compute some numerical examples to demonstrate the usage of the proposed methods. Finally, conclusions are addressed and some future research issues on the test procedure are suggested.

1.1 Review on FMCIA

The original research works on FMCIA can be dated back to 1986 when Fu [1] proposed a rudiment of this method in his study on evaluating the reliability of a consecutive- k -out-of- n system with components dependency. Soon after that, Fu and Hu [2], and Chao and Fu [3,4] applied this method to obtain the reliability

evaluation for different types of system structures. Until 1994, Fu and Koutras [5] formally named this method the FMCIA.

The FMCIA provides an efficient way to perform the reliability analysis, the component importance analysis, failure rate function computations for the system with complex structure that the traditional methods, e.g., the combinatorial method, the total probability formula method, the recursive equation method, and the moment generating function method, are sometimes difficult to do so. Besides the reliability field, many papers that applied the FMCIA have been published to study the statistical process controls, the runs and scans, the hypothesis test, the samplings, the demonstrate test, the management of health, and the DNA sequences analysis.

When evaluating the system reliability, traditional methods are required to enumerate all possible cases to obtain the equation, which is sometimes unable to realize for a large scale of system, especially when the components of the system are nonidentical. Koutras and Alexandrou [6] indicated that the FMCIA can overcome the bottleneck. Since then, a lot of research works on evaluating and analyzing the reliability models for complicated system structure have been accomplished by employing the FMCIA.

Koutras [7] presented a general framework to accommodate the reliability models for various system structures and he obtained some properties for those structures. Chadjiconstantinidis and Koutras [8] studied the reliability importance of components, e.g., the Birnbaum importance, for the wide class of Markov chain embeddable system. Chang et al. [9] expanded the linear (n, f, k) system, which consisted of n components and failed if there exist total- f or consecutive- k failed components, to circular case and presented the reliability equation. Cui et al. [10] modified the two failure criteria model from “or” to “and,” and proposed the $\langle n, f, k \rangle$ model, whose reliability equation is obtained by FMCIA. Guo et al. [11] involved a concept of failure window, i.e., the consecutive failure criterion only applies to component $i, i+1, \dots, j-1, j$, and developed a new reliability models of $(n, f, k(i, j))$ and $\langle n, f, k(i, j) \rangle$. Zhao et al. [12] introduced the concept of “sparsely d connected” and expanded the consecutive- k models to general cases. Zhao and Cui [13] conducted a research to accelerate the calculation speed and reduce the computation complexity for a large scale of system with multiple types of components. Cui et al. [14] published a review article by summarizing the development and application of FMCIA prior to 2010.

The research on system reliability by using this method becomes more diverse. Zhao and Cui [15] replace the binary-state components with multistate components for the k -out-of- n system. Chang and Huang [16] applied the FMCIA to analyze the reliability of system with two-dimensional component alignments. Zhao et al. [17] derived the reliability equation for a system, whose components allocated in matrix shape, with linear connected failure criterion. Zhao et al. [18] obtained the reliability evaluation for two-dimensional system with trinary-state components. Zhao et al. [19] derived the state distribution for a multistate consecutive- k system. By involving the concept of “ l -overlapping between two consecutive- k segments,” Cui et al. [20] proposed the m -consecutive- k , l -out-of- n system model. Shen and Cui [21] applied the FMCIA to calculate the reliability and Birnbaum importance for sparsely connected circular consecutive- k system. Du et al. [22] used the FMCIA to calculate the reliability of a system operating under multiphase missions. Lin

et al. [23] proposed a consecutive- k_r -out-of- n_r :F linear zigzag structure and circular polygon structure.

For other recent research by applying the FMCIA, one can refer to Chang and Wu [24], Fu et al. [25], Koutras and Milienos [26], Fu [27], Fu and Hsieh [28], and Lee [29].

1.2 Illustrate test procedure

The CS (consecutive successes) start-up demonstrated test was proposed by Hahn and Gage [30] and Viveros and Balakrishnan [31]. Then Balakrishnan and Chan [32] modified the previous works and proposed the CSTF (consecutive success total failure) test model. Since then, various types of test models with different acceptance and (or) rejection criterion are developed. For example, Smith and Griffith [33] applied maximum likelihood estimation to calculate the reliability of the test; Eryilmaz and Chakraborti [34], by assuming that the probability of a successful test is a random variable rather than a fixed value, derived the probability distribution of the waiting time of the CSTF model; Gera [35] obtained the average test length (i.e., the expected number of test samples) for the CSTF including the dependent test models. The FMCIA is also employed to calculate the reliability of the test procedure. For instance, Zhao et al. [36] involved the concept of “start-up delay” to the demonstration test procedure for the product; Zhao [37] and Zhao et al. [38] presented a review paper on binary-state and multistate start-up demonstration tests, respectively; Zhao et al. [39] considered the “sparse- d ” concept to the demonstrate test.

In this chapter, the quality test procedure can be described as Fig. 1. We denote the number of tested samples as n . The product is accepted if there are consecutive- k

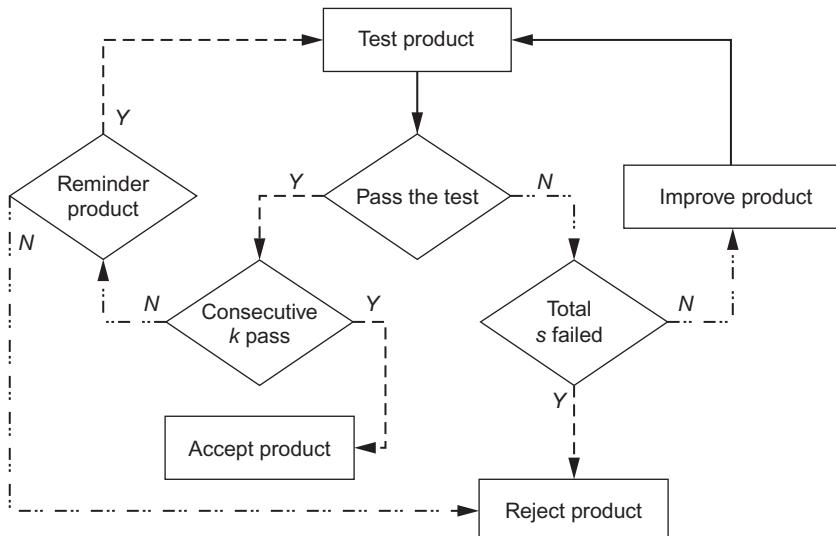


Fig. 1 The illustrative diagram of the quality test procedure.

products passing the test; and the product is rejected if there are total $s = n - k + 1$ products that have been failed the test.

In order to enhance the quality of the product, we carry out an improvement action on each reminder products whenever a failure of test occurs. It is like conducting an adjustment to bad behavior. For example, let p_1 represent the probability that the first test can be passed. Given this round of test failed, we perform an enhancement to the untested samples. If the enhancement action is accomplished, the probability of passing the next round of test increases to p_2 ; otherwise, the probability of success stays p_1 . Thus, we denote p_l representing the probability of passing the test after the $(l - 1)^{th}$ corrective action is successfully carried out, where the probability values of success follow the order of $p_1 < p_2 < \dots < p_l < \dots < p_s$ ($l = 1, 2, \dots, s$). The increment of the probability of passing one test can be in various ways (e.g., linear, exponential, or random), which are presented in the numerical example section.

However, the improvement action on the untested product is influenced by a lot of factors, e.g., the environment fluctuation, the human behaviors, or the cost constrain, so that it may not be accomplished. Therefore, we assume that corrective action can be successfully carried out with probability p . The value of p can be interpreted as the scale parameter for improving the quality of the product. The greater the value is, the greater the product can pass the test procedure. We present the effect of the value of p in the numerical example section.

2 Mathematical description for modeling

By employing FMCIA to analyze the qualification test procedure, we first need to define a Markov chain, $\{X(t), t = 1, 2, \dots, n\}$, with finite state space S :

$$S = \{(i, j), 1 \leq i \leq s = n - k + 1, 1 \leq j \leq k\} \cup S_{N-1} \cup S_N,$$

where state (i, j) indicates that the j^{th} consecutive test is under way, given there has been $(i - 1)$ failed products; states S_{N-1} and S_N indicate the quality test for the product is passed and failed, respectively. Thus, the total number of the states is $N = sk + 2 = (n - k + 1)k + 2$. For example, state $(5, 3)$ means there are four tests failed and the third test is undertaking since the last failed test.

For convenience, we label state (i, j) as number $(i - 1)k + j$, states S_{N-1} as G and S_N as F, so the state space of the Markov chain is $S = \{1, 2, \dots, N - 2, G, F\}$. Therefore, the result of the quality test result can be expressed by the state of the Markov chain, i.e., the products pass the test procedure if there exists t_1 ($1 \leq t_1 \leq n$) so that $X(t_1) = G$; and the products fail the test procedure if there exists t_2 ($1 \leq t_2 \leq n$) so that $X(t_2) = F$ or $X(n) \notin \{G, F\}$.

The state transition of the Markov chain can be categorized into six situations:

- (i) $(i, j) \rightarrow (i, j+1)$ for $i = 1, 2, \dots, s$ and $j = 1, 2, \dots, k - 1$. This situation means that the test is passed. The state transition probability is presented as $P\{(i, j) \rightarrow (i, j+1)\} = p_i$.
- (ii) $(i, j) \rightarrow (i+1, 1)$ for $i = 1, 2, \dots, s - 1$ and $j = 1, 2, \dots, k$. This situation means that the test is failed and the improvement action is accomplished so that the “consecutive- k successful

run” needs to be recounted and the probability for passing the next round of test increases to p_{i+1} . The state transition probability is presented as $P\{(i, j) \rightarrow (i+1, 1)\} = pq_i$.

- (iii) $(i, j) \rightarrow (i, 1)$ for $i = 1, 2, \dots, s-1$ and $j = 1, 2, \dots, k$. This situation means that the test is failed and so does the improvement action. Thus the “consecutive- k successful run” needs to be recounted and the probability for passing the next round of test stays p_i . The state transition probability is presented as $P\{(i, j) \rightarrow (i, 1)\} = qq_i$.
- (iv) $(i, k) \rightarrow G$ for $i = 1, 2, \dots, s$. This situation means that the product passes the test procedure and is accepted because the consecutive- k successful test criterion is achieved. The state transition probability is presented as $P\{(i, k) \rightarrow G\} = p_i$.
- (v) $(s, j) \rightarrow F$ for $j = 1, 2, \dots, k$. This situation means that the product fails the test procedure and is rejected, as the total number of failed test exceeds the limit of s . The state transition probability is presented as $P\{(s, j) \rightarrow F\} = q_s$.
- (vi) $F \rightarrow F$ and $G \rightarrow G$. F and G are two absorbing states. Thus the state transition probability is presented as $P\{F \rightarrow F\} = P\{G \rightarrow G\} = 1$.

To sum up, the transition probability matrix of the defined Markov chain is presented as a block matrix $\Lambda(t)$ and the transition diagram is shown in Fig. 2.

$$\Lambda(t) = \begin{pmatrix} (C_{k \times k}(i, j))_{s \times s} & D_{k \times 2}(1) \\ \vdots & D_{k \times 2}(s) \\ \mathbf{0}_{2 \times sk} & I_{2 \times 2} \end{pmatrix}_{|sk+2| \times |sk+2|},$$

where $C_{k \times k}(i, j) = \mathbf{0}_{k \times k}$ for $i \neq j$ and $i \neq j+1$ ($i = 1, 2, \dots, s-1$),

$$C_{k \times k}(i, i) = \begin{pmatrix} qq_i & p_i & 0 & \cdots & 0 \\ qq_i & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & p_i \\ qq_i & 0 & \cdots & \cdots & 0 \end{pmatrix}_{k \times k}, \quad C_{k \times k}(i, i+1) = \begin{pmatrix} pq_i & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ pq_i & 0 & \cdots & 0 \end{pmatrix}_{k \times k},$$

$$C_{k \times k}(s, s) = \begin{pmatrix} 0 & p_s & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & \ddots & p_s \\ \mathbf{0} & & & & 0 \end{pmatrix}_{k \times k}, \quad D_{k \times 2}(i) = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ p_i & 0 \end{pmatrix}, \quad (i = 1, 2, \dots, s-1),$$

$$D_{k \times 2}(s) = \begin{pmatrix} 0 & q_s \\ \vdots & \vdots \\ 0 & q_s \\ p_s & q_s \end{pmatrix}$$

$I_{2 \times 2}$ is a unit matrix, $\mathbf{0}$ and $\mathbf{0}_{k \times k}$ are zero matrixes.

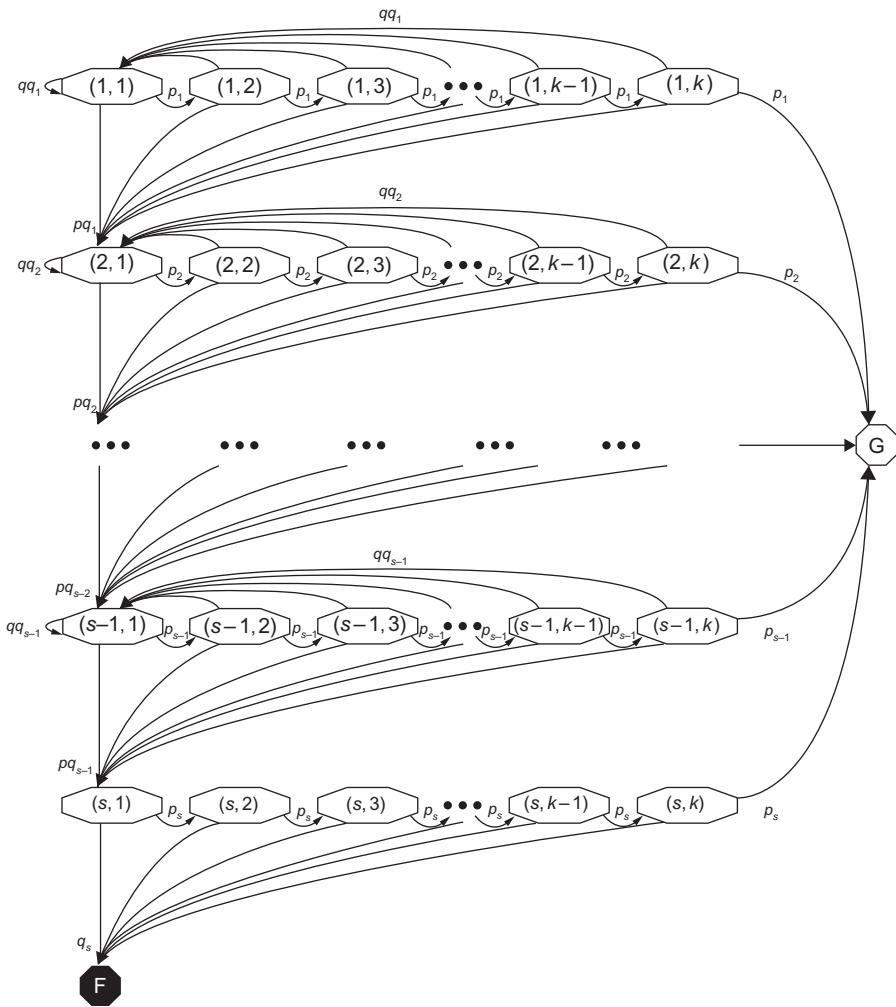


Fig. 2 The state transition diagram of $\{X(t), t = 1, 2, \dots, n\}$.

3 Quantitative analysis for the test procedure

For a quality test procedure, there are some indexes such as the probability that the product is accepted (or rejected), the expected number of tested samples for stopping the test regardless of acceptance or rejection, and expected number of improvements action taken during the test. In this section we derived the equations for obtaining those indexes.

3.1 Probability of acceptance

As is declared, the product can be accepted only when there are consecutive- k successful tests before all n samples are tested. Therefore, the probability of acceptance is the probability that the Markov chain finally stays in state G. Let the initial state probability vector of the defined Markov chains be $\boldsymbol{\pi}_0 = (1, 0, \dots, 0)_{1 \times |sk+2|}$, because at the very beginning, there are zero failed test and the first test is undergoing, i.e., the Markov chain stays in state (1, 1). Then in terms of Chapman–Kolmogorov equation, we can obtain the probability that the product passes the test procedure and is accepted by Eq. (1), where the function of $\mathbf{U}_G = (0, \dots, 0, 1, 0)_{1 \times |sk+2|}$ is to extract the probability value so that the product is accepted; and the symbol T is a matrix transpose operator.

$$P\{\text{accept}\} = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^n \mathbf{U}_G^T. \quad (1)$$

$\boldsymbol{\Lambda}(t)$ is the one-step transition probability matrix of the defined Markov chain. Thus, $\boldsymbol{\Lambda}(t)^n$ is the n -step transition probability matrix of the Markov chain. Given the initial state vector $\boldsymbol{\pi}_0$, we can obtain the probability value that the Markov chain is in each state after n -step transition by $\boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^n$. Finally, we extract the probability value that the test procedure is passed, i.e., the Markov chain stays at state G, by multiplying \mathbf{U}_G^T .

Accordingly, the probability for rejecting the product is calculated by Eq. (2), where $\mathbf{U}_{\text{reject}} = (1, \dots, 1, 0, 1)_{1 \times |sk+2|}$.

$$P\{\text{reject}\} = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^n \mathbf{U}_{\text{reject}}^T. \quad (2)$$

3.2 Expected number of tests

The test procedure is terminated under the following three cases: (i) the product is accepted as there have been consecutive- k successful tests occurred, i.e., the Markov chain arrives at state G after l steps; (ii) the product is rejected as there have been total s failed tests appeared, i.e., the Markov chain arrives at state F after several steps of transitions; (iii) the product is rejected as all n products have been tested but neither of the criteria for acceptance and rejection has been accomplished, i.e., the Markov chain arrives at state $(i, 1)$ after l ($l \geq n - k + 1$) steps.

For the first case, let $P_G(l)$ represent the probability of the Markov chain arrives at state G after l ($l \geq k$) steps of transitions, which is calculated by Eq. (3). We denote N_G as the number of tested samples when the test procedure is passed. Then, the expected value of N_G is calculated by Eq. (4), where the term $[P_G(l) - P_G(l-1)]$ is the probability that the Markov chain arrives at state G for the first time after l ($l = k, k+1, \dots, n$) steps.

$$P_G(l) = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^l \mathbf{U}_G^T. \quad (3)$$

$$\begin{aligned} E(N_G) &= kP_G(k) + \sum_{l=k+1}^n l[P_G(l) - P_G(l-1)] \\ &= nP_G(n) - \sum_{l=k}^{n-1} P_G(l). \end{aligned} \quad (4)$$

For the second case, similarly, we denote $P_F(l)$ representing the probability of the Markov chain that arrives at state F after l ($l \geq s$) tests, which can be calculated by Eq. (5), where $\mathbf{U}_F = (0, \dots, 0, 1)_{1 \times |sk+2|}$. Let N_F be the number of tested samples when the test procedure is rejected due to the fact that there is at least s failed tests occurred. Thus, The expected value of N_F is calculated by Eq. (6).

$$P_F(l) = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^l \mathbf{U}_F^T. \quad (5)$$

$$E(N_F) = nP_F(n) - \sum_{l=s}^{n-1} P_F(l). \quad (6)$$

For the third case, since the Markov chain arrives at state $(i, 1)$ after l ($l \geq n-k+1$) steps, we have $i = \lceil \frac{n-k}{k-1} \rceil$. We denote $P_F((i, 1), l)$ as the probability of the Markov chain arriving at state $(i, 1)$, after l tests, which is calculated by Eq. (7), where the function of $\mathbf{U}_{(i, 1)} = \left(0, \dots, 0, \underset{1 \times |sk+2|}{\underset{(i-1)k+1}1}, 0, \dots, 0\right)$ is to extract the probability value that the Markov chain arrived in state $(i, 1)$.

$$P_F((i, 1), l) = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^l \mathbf{U}_{(i, 1)}^T. \quad (7)$$

Let $N_F(i, 1)$ represent the number of tested samples when the Markov chain arrives at state $(i, 1)$ after l tests. Thus, the expected value of $N_F(i, 1)$ is calculated by Eq. (8), where $\lceil x \rceil$ is the smallest integer which is no less than x .

$$E(N_F(i, 1)) = \sum_{i=\lceil \frac{n-1}{k-1} \rceil}^s \sum_{l=s}^n l P_F((i, 1), l). \quad (8)$$

Since the three cases are disjoint, based on the results for the three cases, the expected number of tested samples equals the summation of those three cases and is obtained by Eq. (9).

$$E(N_{\text{test}}) = E(N_G) + E(N_F) + E(N_F(i, 1)). \quad (9)$$

3.3 Expected number of improvement actions

The number of improvement actions is related to the number of the failed test. We still follow the three cases that lead the termination of the test procedure, as illustrated already.

For the first case, as the test procedure stops by accepting the product, it means that Markov chain must arrive at state (i, k) for $i = 1, 2, \dots, s$ after $(l - 1)$ steps, whose probability is presented by Eq. (10), where $\mathbf{U}_{(i, k)} = \left(0, \dots, 0, \overset{(ik)\text{th}}{1}, 0, \dots, 0 \right)_{1 \times |sk+2|}$, and the l th test must be successful with probability p_i . We denote by \tilde{N}_G the number of the improvements with test procedure passing. Then the expected value of \tilde{N}_G is calculated by Eq. (11) because the number of successful carried-out improvement actions is $i - 1$.

$$P((i, k), l - 1) = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^{l-1} \mathbf{U}_{(i, k)}^T. \quad (10)$$

$$E(\tilde{N}_G) = \sum_{i=1}^s \left[(i-1) \sum_{l=k}^n P((i, k), l - 1) p_i \right]. \quad (11)$$

For the second case, similarly, as the test procedure stops by rejecting the product before all samples are tested, the Markov chain must arrive at state (s, j) for $j = 1, 2, \dots, k$ after $(l - 1)$ steps and the l th test must be failed with probability q_s . The probability that the Markov chain stays at (s, j) after $(l - 1)$ steps is calculated

by Eq. (12), where $\mathbf{U}_{(s, j)} = \left(0, \dots, 0, \overset{((s-1)k+j)\text{th}}{1}, 0, \dots, 0 \right)_{1 \times |sk+2|}$. Let \tilde{N}_F represent

the number of the improvements with test procedure fails. The expected value of \tilde{N}_F is calculated by Eq. (13), because, in this case, the number of the successfully carried-out improvement actions is $s - 1$.

$$P((s, j), l - 1) = \boldsymbol{\pi}_0 \boldsymbol{\Lambda}(t)^{l-1} \mathbf{U}_{(s, j)}^T. \quad (12)$$

$$E(\tilde{N}_F) = (s-1) \sum_{j=1}^k \sum_{l=s}^n [P((s, j), l - 1) q_s]. \quad (13)$$

For the third case, the test procedure is failed due to the n products that are all tested before the criteria of acceptance or rejection is met. Thus the Markov chain must arrive at state $(i, 1)$ after l steps. The number of improvements of this case is denoted as $\tilde{N}_F(i, 1)$. Thus, the expected value of $\tilde{N}_F(i, 1)$ is presented as Eq. (14), by employing Eq. (7), because the number of the successfully carried-out improvement action is $i - 1$.

$$E(\tilde{N}_F(i, 1)) = \sum_{i=1}^s \left[(i-1) \sum_{l=s}^n P_F((i, 1), l) \right]. \quad (14)$$

To sum up, the expected number of successful improvements can be obtained by Eq. (15).

$$E(\tilde{N}_{\text{improvement}}) = E(\tilde{N}_G) + E(\tilde{N}_F) + E(\tilde{N}_F(i, 1)). \quad (15)$$

4 Numerical examples

In this section, we apply five numerical examples to illustrate how to use the derived equations to calculate the proposed indexes and analyze the test procedure. The first example explains the effect of the successful probability of improvement action on the indexes. The second example reveals the effect of different initial quality state of the product and improvement scenarios on the indexes. The third example compares the effect of the successful probability of improvement action with the same level of initial quality and different improvement scenarios on the indexes. The fourth example exhibits the effect of the total number of tested samples (e.g., n) and the acceptance criteria (e.g., consecutive- k) on the indexes.

Example 1

We assume the total number of tested samples as $n = 6$. The product is accepted if there are consecutive $k = 3$ tests passed, or is rejected if there are total $s = n - k + 1 = 4$ tests failed. The improvement action is performed whenever a test is failed. In this example, we assume that the improvement can increase the probability of passing the next round of test in a linear way, i.e., $p_1 = 0.8$, $p_2 = 0.85$, $p_3 = 0.9$, and $p_4 = 0.95$. The probability that the improvement can be accomplished, p , varies from 0 to 1. When $p = 0$, it can be interpreted as there is no improvement action after the failed test. When $p = 1$, it means the remainder samples are always enhanced after each failed test. We can obtain the state transition probability matrix as $\Lambda(t)$.

$$\Lambda(t) = \left(\begin{array}{ccc|ccc|c|cc} qq_1 & p_1 & 0 & pq_1 & 0 & 0 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & 0 & 0 \\ qq_1 & 0 & p_1 & pq_1 & 0 & 0 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & 0 & 0 \\ qq_1 & 0 & 0 & pq_1 & 0 & 0 & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 3} & p_1 & 0 \\ \hline & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ \hline & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & p_2 & 0 \\ \hline & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & p_3 & 0 \\ \hline & & & & & & & & 0 & p_4 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & 0 & q_4 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & p_4 & q_4 \\ \hline & & & & & & & & 1 & 0 \\ & & & & & & & & 0 & 1 \\ \hline \mathbf{0}_{2 \times 3} & 14 & 14 \end{array} \right)$$

The plots of the probability of acceptance, the expected number of tests conducted, and the expected number of improvements according to different p are presented in Figs. 3 and 4, respectively.

The results reveal that the greater the p is, the more the chance that the product is accepted, as the improvement action make the product toward good quality. However,

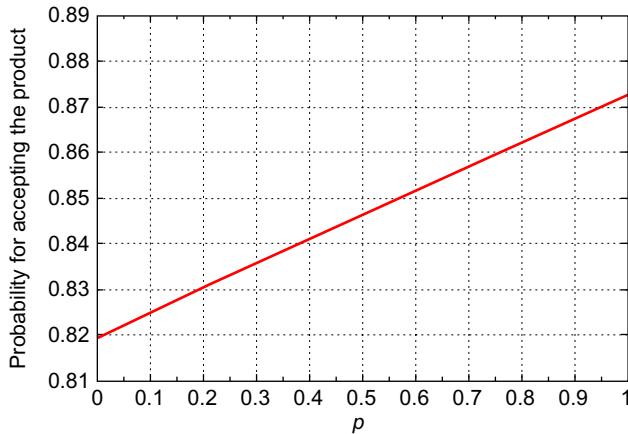


Fig. 3 Probability for accepting the product according to p .

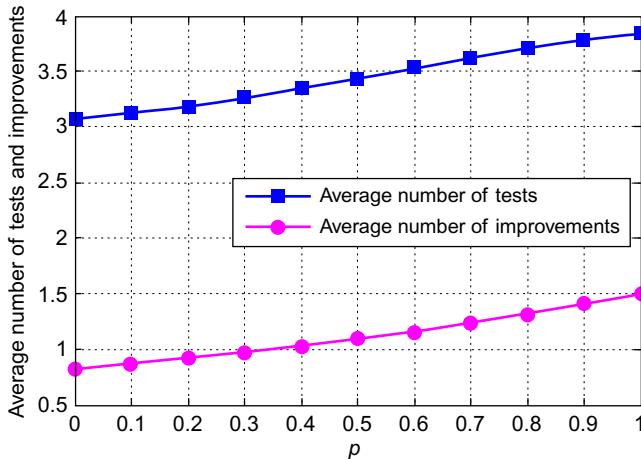


Fig. 4 Expected number of tests and improvements according to p .

the expected number of tests and improvement actions are increased, which leads the cost for the test procedure grows. These indexes can be used for optimizing the test plan in further research.

Example 2

We use the same parameters for the test procedure, i.e., $n = 6$, $k = 3$, and $s = 4$; and we assume the probability that the improvement is successfully made equal to $p = 0.8$. The quality of the product is enhanced after each improvements action, i.e., the probability for passing the next round of test is increased, $p_{l+1} > p_l$.

In some situations, as the improvement action continues, the increase of p_l may converge. It is because as the quality of the product is improved, it is hard to make

any further progress on the product, which is in accordance with our intuitive feeling. Therefore, in this example, we assume the probability for passing the next round of test increases in an exponential way, which is presented by Eq. (16), where $\lambda \in (0, 1)$. The probability values of p_l ($l = 1, 2, \dots, s$) for some different value of λ are listed in Table 1.

$$p_l = 1 - \lambda e^{-\lambda(l-1)}. \quad (16)$$

Then, the probability values that the product can be accepted for different improvement scenarios are described in Fig. 5, which is decreasing first and then increasing, e.g., $P\{\text{accept}|\lambda=0.6\} < P\{\text{accept}|\lambda=0.7\} < P\{\text{accept}|\lambda=0.8\}$. The result shows that the different levels of the initial quality of the product and the improvement rate (i.e., λ) have different effect on the test procedure.

Table 1 Probability values of p_l for some different value of λ

	p_1	p_2	p_3	p_4
$\lambda = 0.1$	0.9000	0.9095	0.9181	0.9259
$\lambda = 0.2$	0.8000	0.8363	0.8659	0.8902
$\lambda = 0.3$	0.7000	0.7778	0.8354	0.8780
$\lambda = 0.4$	0.6000	0.7319	0.8203	0.8795
$\lambda = 0.5$	0.5000	0.6967	0.8161	0.8884
$\lambda = 0.6$	0.4000	0.6707	0.8193	0.9008
$\lambda = 0.7$	0.3000	0.6524	0.8274	0.9143
$\lambda = 0.8$	0.2000	0.6405	0.8385	0.9274
$\lambda = 0.9$	0.1000	0.6341	0.8512	0.9395

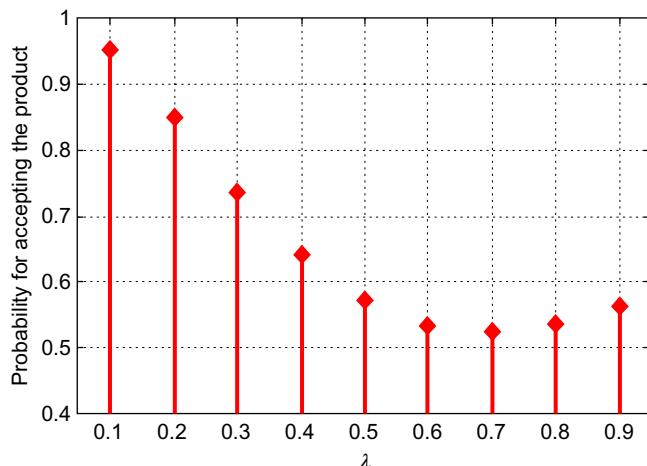


Fig. 5 Probability for accepting the product according to λ .

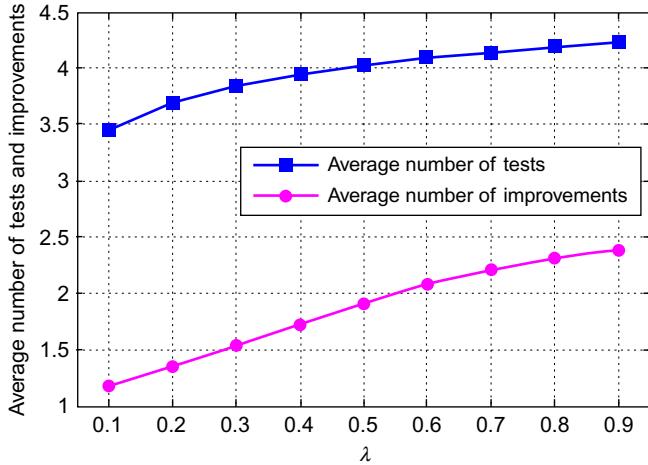


Fig. 6 Expected number of tests and improvements according to λ .

As is presented in Fig. 6, the expected number of tests and improvements are both increasing. The result shows that the bad initial quality of the product will lead the expected number of the tests and improvement actions to increase. Thus, the test procedure lasts long and costs more. But as the rate growth improves, those expected values converge.

Example 3

In this example, we still fix the parameters of the test procedure, i.e., $n = 6$, $k = 3$, and $s = 4$; and we assume the probability that the improvement action, p , varies from zero to one. In this example, we assume the initial quality of the product to be the same with $p_1 = 0.6$. We assign different improvement scenarios, which are presented in Table 2, whenever a failure of test occurs.

In each scenario, we change only one probability value of the probability for passing the next round test. To be specific, we set scenario as control group; Scenarios 2 and 3 represent the different levels of the first improvement action; Scenarios 4 and 5

Table 2 Probability values of p_l for some different improvement scenarios

	p_1	p_2	p_3	p_4
Scenario 1	0.6000	0.7000	0.8000	0.9000
Scenario 2	0.6000	0.7500	0.8000	0.9000
Scenario 3	0.6000	0.6500	0.8000	0.9000
Scenario 4	0.6000	0.7000	0.8500	0.9000
Scenario 5	0.6000	0.7000	0.7500	0.9000
Scenario 6	0.6000	0.7000	0.8000	0.9500
Scenario 7	0.6000	0.7000	0.8000	0.8500

represent the different levels of the second improvement action; Scenarios 6 and 7 represent the different levels of the third improvement action.

The comparisons of the probability that the product can pass the test procedure for different scenarios are presented in Fig. 7A–C. Indeed, the greater the probability for implementing the improvement actions is, the higher the probability of the acceptance. Besides, the results also reveal that the improvement actions should be taken earlier so that the probability of acceptance can grow faster. For example, by comparing the results of Scenarios 2, 4, and 6 with Scenario 1, it shows that the earlier improvement action can lead the probability of acceptance increase more. The same conclusion can be obtained by comparing Scenarios 3, 5, and 7 with Scenario 1. The results show that if the earlier improvement action is taken with lower extent, the probability of acceptance decreases more evidently.

The comparisons of the expected number of tests for different scenarios are obtained and presented in Fig. 8A–C. The result shows that when the probability for successfully implementing the improvement actions is lower (e.g., $p < 0.7$), the expected number of tests is almost the same for different scenarios; when the probability value of p is greater, the results deviate from that of the control group.

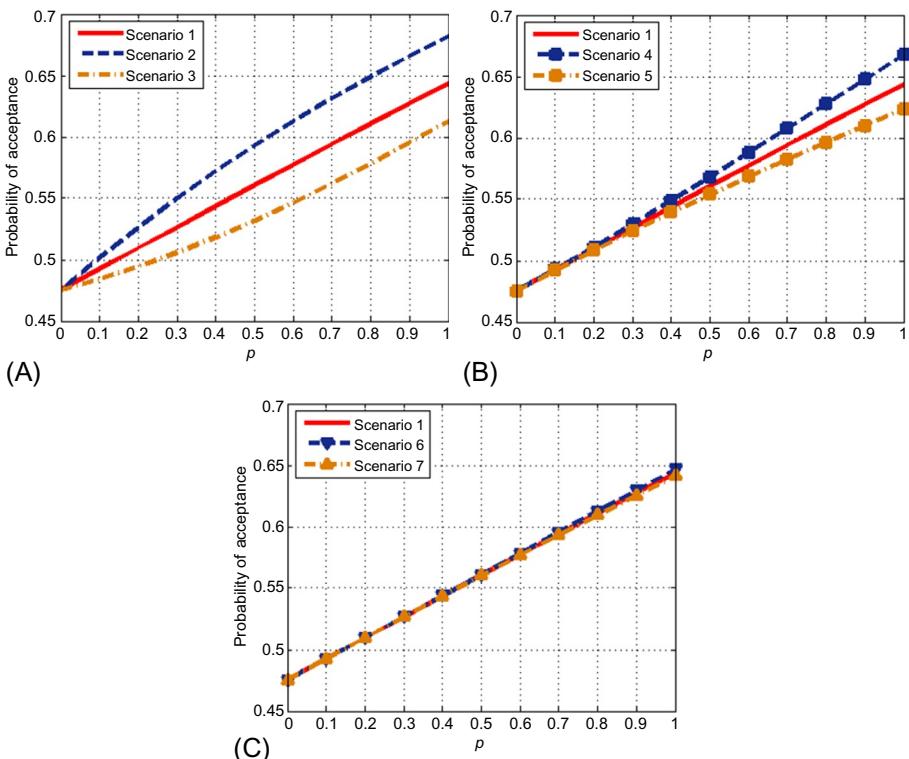


Fig. 7 Comparison of probability of acceptance for different scenarios.

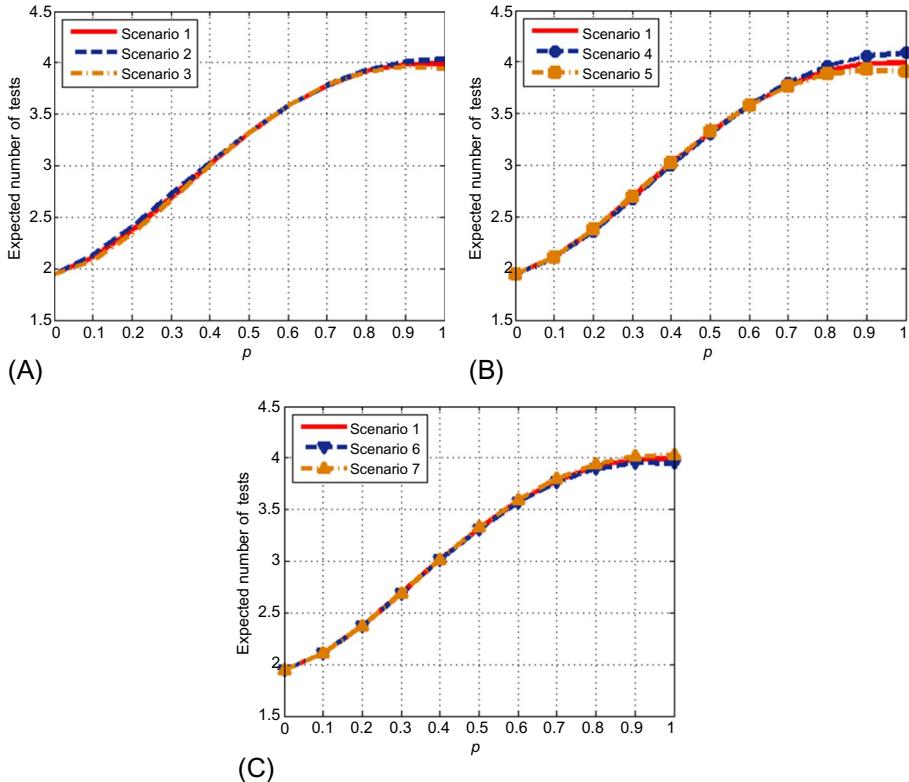


Fig. 8 Comparison of expected number of tests for different scenarios.

The comparisons of the expected number of improvement actions for different scenarios are depicted in Fig. 9A–C. The comparison results show that a better improvement can reduce the expected number of improvement actions. For example in Fig. 9A, Scenario 2 takes a better enhancement on the first improvement action than Scenario 1 (i.e., p_2 in Scenario 2 is greater than that in Scenario 1); but Scenario 3 takes a worse enhancement on the first improvement action than Scenario 1 (i.e., p_2 in Scenario 3 is lower than that in Scenario 1). So the expected number of improvement actions is in the order of $E(\tilde{N}_{\text{improvement}}^{\text{Scenario 2}}) < E(\tilde{N}_{\text{improvement}}^{\text{Scenario 1}}) < E(\tilde{N}_{\text{improvement}}^{\text{Scenario 3}})$ for these three scenarios. The similar results can be concluded for Fig. 9B and C. The explanations are omitted here.

Example 4

The parameters of the test procedure (i.e., the number of total tested samples n and the acceptance criteria: consecutive- k) affect the process of the test. For example, the greater the value of k is, the lower the probability of passing the test procedure, because the criteria for accepting the product becomes strict. We apply the proposed method to a numerical example to present the influence caused by those parameters.

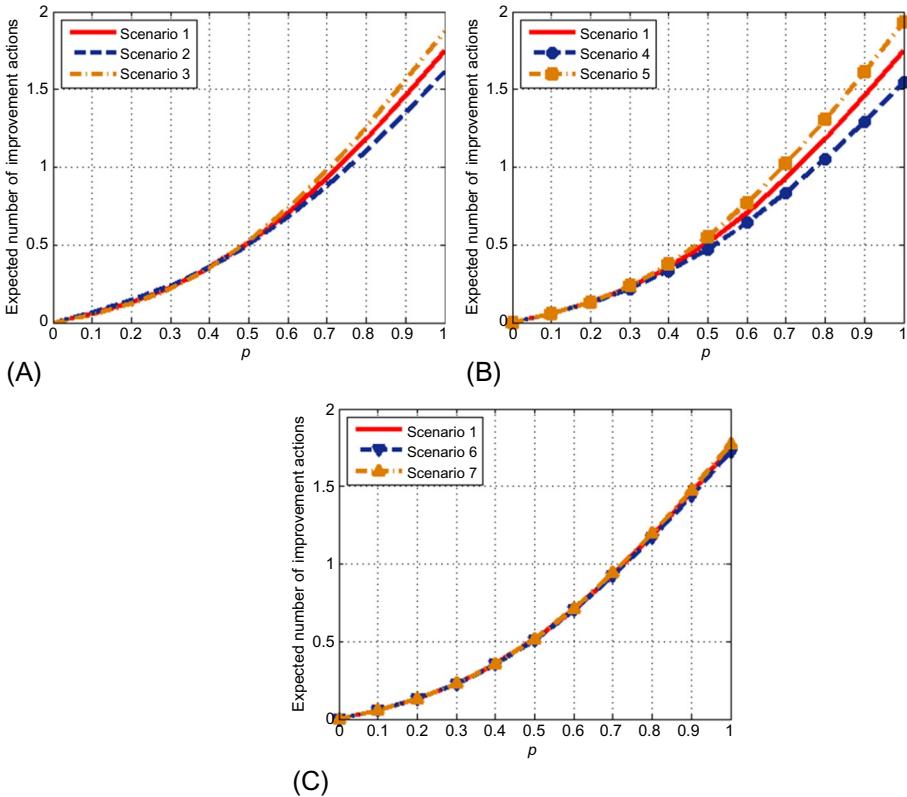


Fig. 9 Comparison of expected number of improvement actions for different scenarios.

We assume that the probability value of passing the next round of test, p_l , is improved randomly, which is given in [Table 3](#). We set the probability for accomplishing the improvement to be $p = 0.8$.

We assume $2 \leq k \leq 10$ and $15 \leq n \leq 30$. The probability of accepting the product is presented in [Fig. 10](#). The result derives the following conclusions. For a fixed value of k ,

Table 3 Probability for passing the next round of test after improvements

$p_1 \rightarrow p_6$	$p_7 \rightarrow p_{12}$	$p_{13} \rightarrow p_{18}$	$p_{19} \rightarrow p_{24}$	$p_{25} \rightarrow p_{29}$
0.5159	0.5934	0.7228	0.8474	0.9117
0.5172	0.6385	0.7449	0.8530	0.9246
0.5179	0.6585	0.8232	0.8716	0.9670
0.5231	0.6908	0.8277	0.8789	0.9751
0.5486	0.6961	0.8279	0.8828	0.9797
0.5856	0.7194	0.8394	0.8976	—

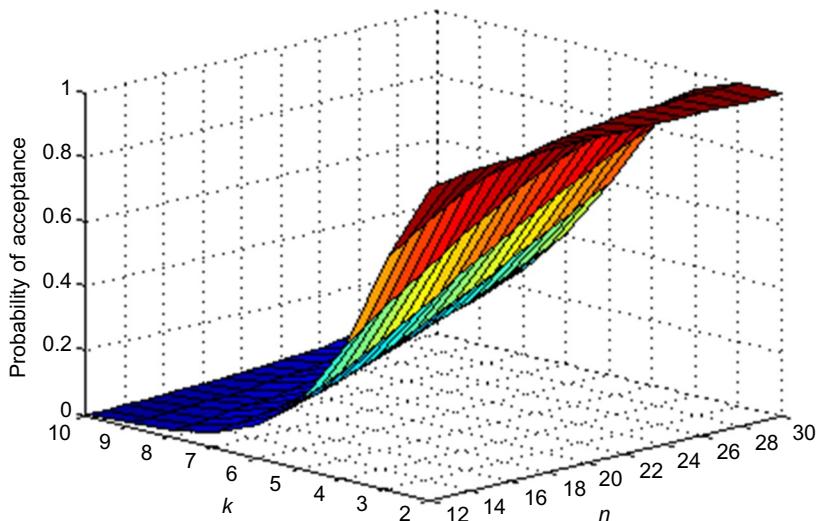


Fig. 10 Probability of acceptance according to k and n .

the greater the value of n is, the higher the probability for accepting the product is. It is because the acceptance criterion becomes relatively slack so that the product is easy to pass the test procedure. On the contrary, for a fixed value of n , the greater the value of k is, the lower the probability that the product that can pass the test procedure is. It is because the acceptance criterion becomes relative tight so that the test procedure tends to be failed.

The average number of tests that need to be conducted for the combination of k and n are presented in Fig. 11. The interpretations for the result can be as follows: For the very slack criterion, e.g., $k = 2$, the test procedure is very likely to be passed within fewer tests; for the slack criterion, e.g., $k = 4$ and $n = 30$, as the probability of acceptance is high, the test procedure is likely to be passed after several time tests; for the high criterion, e.g., $k = 10$ and $n = 12$, as the probability of acceptance is low, the test is likely to be failed with a number of trials.

The expected number of the improvement actions that needs to be performed for different combinations of k and n is presented in Fig. 12. For the slack criterion, as the test procedure is more likely to be passed, the number of improvement action tends to be small. But for the tight criterion, the number of improvement is expected to be greater, because the test is inclined to be failed so that enhancement action is more likely to be occurred.

To sum up, in this example, we suggest that the parameter of the test plan is set to be $k = 4$ and $n = 14$ around because for this combination, the probability for passing the test procedure is high; the expected number of test need to be conducted is relative small; and the expected number of improvement action, which is cost consumed, that is required to be performed is very small.

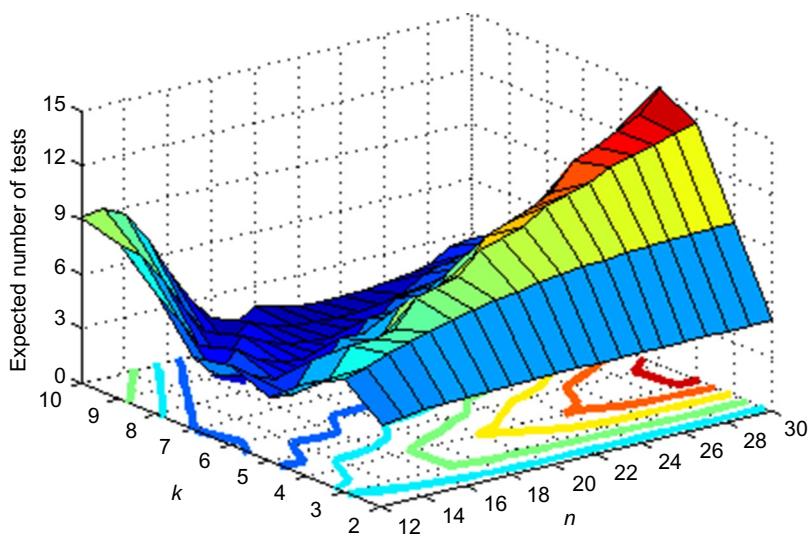


Fig. 11 Expected number of tests according to k and n .

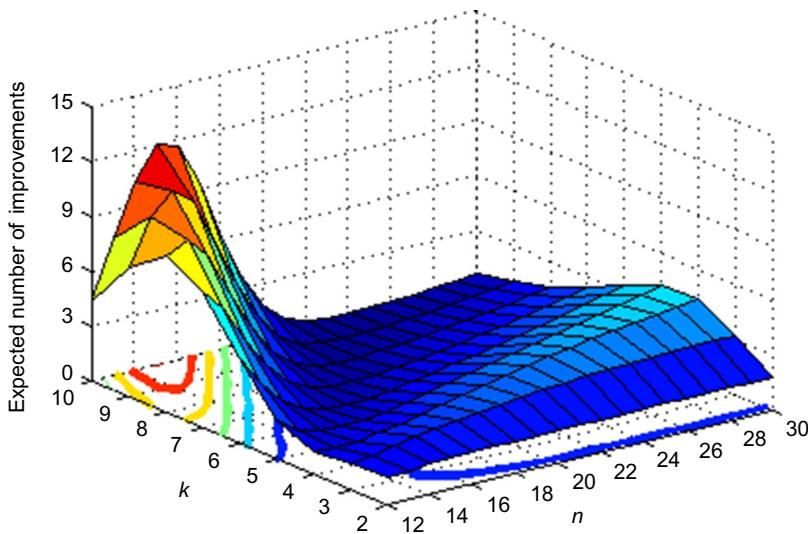


Fig. 12 Expected number of improvements according to k and n .

5 Conclusions

The qualification test procedure subjected to the consecutive- k acceptance criterion is studied in this chapter by using the technique of FMCIA. Some indexes, such as the probability of accepting the product, the expected number of tested samples, and the

expected number of improvements taken after each failed test during the procedure, are proposed. We present some numerical examples as an illustration of the proposed methods.

In real practice, besides the consecutive- k criterion, there are still a lot of factors that need to be considered for qualification tests. For example, as the test procedure is a cost-related process, the optimal decision to maximize the probability of success and to minimize the cost spent is always an issue that needs to be trade-off. Thus, if we change the testing rules, new models are evolved. For those models or problems, we believe that the FMCIA is an efficient way to solve them.

Acknowledgments

The authors are supported by NSFC under project 71631001. The authors would like to express their thanks to it.

References

- [1] J.C. Fu, Reliability of consecutive- k -out-of- n :F systems with $(k-1)$ -step Markov dependence, *IEEE Trans. Reliab.* 77 (5) (1986) 602–606.
- [2] J.C. Fu, B. Hu, On reliability of a large consecutive- k -out-of- n :F system with $(k-1)$ -step Markov dependence, *IEEE Trans. Reliab.* 36 (1) (1987) 75–77.
- [3] M.T. Chao, J.C. Fu, A limit theorem of certain repairable systems, *Ann. Inst. Stat. Math.* 41 (4) (1989) 809–818.
- [4] M. Chao, J. Fu, The reliability of a large series system under Markov structure, *Adv. Appl. Probab.* 23 (23) (1991) 894–908.
- [5] J.C. Fu, M.V. Koutras, Distribution theory of runs: a Markov chain approach, *J. Am. Stat. Assoc.* 88 (427) (1994) 1050–1058.
- [6] M.V. Koutras, V.A. Alexandrou, Runs, scans and urn model distributions: a unified Markov chain approach, *Ann. Inst. Stat. Math.* 47 (4) (1995) 743–766.
- [7] M.V. Koutras, On a Markov chain approach for the study of reliability structures, *J. Appl. Probab.* 33 (2) (1996) 357–367.
- [8] S. Chadjiconstantinidis, M.V. Koutras, Measures of component importance for Markov chain imbeddable reliability structures, *Nav. Res. Logist.* 46 (6) (1999) 613–639.
- [9] G.J. Chang, L.R. Cui, F.K. Hwang, Reliabilities for (n, f, k) systems, *Stat. Probab. Lett.* 43 (3) (1999) 237–242.
- [10] L.R. Cui, et al., On the dual reliability systems of (n, f, k) and $\langle n, f, k \rangle$, *Stat. Probab. Lett.* 76 (11) (2006) 1081–1088.
- [11] Y.L. Guo, et al., Reliabilities for $(n, f, k(i, h))$ and $[n, f, k(i, j)]$ systems, *Commun. Stat.-Theor. Meth.* 35 (10) (2006) 1779–1789.
- [12] X. Zhao, L.R. Cui, W. Kuo, Reliability for sparsely connected consecutive- k systems, *IEEE Trans. Reliab.* 56 (3) (2007) 516–524.
- [13] X. Zhao, L.R. Cui, On the accelerated scan finite Markov chain imbedding approach, *IEEE Trans. Reliab.* 58 (2) (2009) 383–388.
- [14] L.R. Cui, Y. Xu, X.A. Zhao, Developments and applications of the finite Markov chain imbedding approach in reliability, *IEEE Trans. Reliab.* 59 (4) (2010) 685–690.
- [15] X. Zhao, L.R. Cui, Reliability evaluation of generalised multi-state k -out-of- n systems based on FMCI approach, *Int. J. Syst. Sci.* 41 (12) (2010) 1437–1443.

- [16] Y.M. Chang, T.H. Huang, Reliability of a 2-dimensional k -within-consecutive- $r \times s$ -out-of- $m \times n$:F system using finite Markov chains, *IEEE Trans. Reliab.* 59 (4) (2010) 725–733.
- [17] X. Zhao, et al., Exact reliability of a linear connected- (r, s) -out-of- (m, n) :F system, *IEEE Trans. Reliab.* 60 (3) (2011) 689–698.
- [18] X. Zhao, W. Zhao, W.J. Xie, Two-dimensional linear connected- k system with trinary states and its reliability, *J. Syst. Eng. Electron.* 22 (5) (2011) 866–870.
- [19] X. Zhao, Y. Xu, F.Y. Liu, State distributions of multi-state consecutive- k systems, *IEEE Trans. Reliab.* 61 (2) (2012) 274–281.
- [20] L.R. Cui, C. Lin, S.J. Du, m -Consecutive- k , l -out-of- n systems, *IEEE Trans. Reliab.* 64 (1) (2015) 386–393.
- [21] J.Y. Shen, L.R. Cui, Reliability and Birnbaum importance for sparsely connected circular consecutive- k systems, *IEEE Trans. Reliab.* 64 (4) (2015) 1140–1157.
- [22] S.J. Du, C. Lin, L.R. Cui, Reliabilities of a single-unit system with multi-phased missions, *Commun. Stat.-Theor. Meth.* 45 (9) (2016) 2524–2537.
- [23] C. Lin, et al., Reliability modeling on consecutive- k_r -out-of- n_r :F linear zigzag structure and circular polygon structure, *IEEE Trans. Reliab.* 65 (3) (2016) 1509–1521.
- [24] Y.M. Chang, T.L. Wu, On average run lengths of control charts for autocorrelated processes, *Methodol. Comput. Appl. Probab.* 13 (2) (2011) 419–431.
- [25] J.C. Fu, T.L. Wu, W.Y.W. Lou, Continuous, discrete, and conditional scan statistics, *J. Appl. Probab.* 49 (1) (2012) 199–209.
- [26] M.V. Koutras, F.S. Milienos, Exact and asymptotic results for pattern waiting times, *J. Stat. Plan. Inference* 142 (6) (2012) 1464–1479.
- [27] J.C. Fu, On the distribution of the number of occurrences of an order-preserving pattern of length three in a random permutation, *Methodol. Comput. Appl. Probab.* 14 (3) (2012) 831–842.
- [28] J.C. Fu, Y.F. Hsieh, On the distribution of the length of the longest increasing subsequence in a random permutation, *Methodol. Comput. Appl. Probab.* 17 (2) (2015) 489–496.
- [29] W.C. Lee, Power of discrete scan statistics: a finite Markov chain imbedding approach, *Methodol. Comput. Appl. Probab.* 17 (3) (2015) 833–841.
- [30] G.J. Hahn, J.B. Gage, Evaluation of a start-up demonstration test, *PLoS One* 10 (4) (1983) 370–377.
- [31] R. Viveros, N. Balakrishnan, Statistical-inference from start-up demonstration test data, *J. Qual. Technol.* 25 (2) (1993) 119–130.
- [32] N. Balakrishnan, P.S. Chan, Start-up demonstration tests with rejection of units upon observing d failures, *Ann. Inst. Stat. Math.* 52 (1) (2000) 184–196.
- [33] M.L.D. Smith, W.S. Griffith, Start-up demonstration tests based on consecutive successes and total failures, *J. Qual. Technol.* 37 (3) (2005) 186–198.
- [34] S. Eryilmaz, S. Chakraborti, On start-up demonstration tests under exchangeability, *IEEE Trans. Reliab.* 57 (4) (2008) 627–632.
- [35] A.E. Gera, A start-up demonstration procedure involving dependent tests, *Stat. Probab. Lett.* 83 (10) (2013) 2191–2196.
- [36] X. Zhao, Start-up demonstration tests for products with start-up delay, *Qual. Technol. Quant. Manage.* 10 (3) (2013) 329–338.
- [37] X. Zhao, On generalized start-up demonstration tests, *Ann. Oper. Res.* 212 (1) (2014) 225–239.
- [38] X. Zhao, et al., On generalized multi-state start-up demonstration tests, *Appl. Stoch. Model. Bus. Ind.* 31 (3) (2015) 325–338.
- [39] X. Zhao, X.Y. Wang, G. Sun, Start-up demonstration tests with sparse connection, *Eur. J. Oper. Res.* 243 (3) (2015) 865–873.

Quantitative security evaluation of an intrusion tolerant system

2

H. Okamura, T. Dohi
Hiroshima University, Higashi-Hiroshima, Japan

1 Introduction

Because the Internet is highly vulnerable to Internet epidemics, many attacking events compromise a huge number of host computers rapidly, and cause denial-of-service (DoS) around the Internet. Such epidemics result in extensive widespread damage, costing billions of dollars; and countering the propagating worms sufficiently fast becomes an emergency issue for Internet security. Although traditional security approaches, which may be categorized by their *intrusion detection approaches*, establish proactive barriers like a firewall, unfortunately, the efficiency of a single barrier is not still enough to prevent attack from sophisticated new skills by malicious attackers. As a result, the number of network attack incidents is increasing each day. In contrast, to pursue the near impossibility of a perfect barrier unit, the concept of *intrusion tolerance* is becoming more popular. An intrusion tolerant system can avoid severe security failures caused by intrusion and/or attack, and can provide the intended services to users in a timely manner, even under attack. This technique is inspired from traditional techniques commonly used for tolerating accidental faults in hardware and/or software systems, and can provide the system dependability which is defined as a property of a computer-based system, such that reliance can justifiably be placed on the service it delivers [1]. So far, most efforts in security have been focused on specification, design, and implementation issues.

In fact, several implementation techniques for intrusion tolerance at the architecture level have been developed for real computer-based systems such as the distributed systems [2], database systems [3], middleware [4, 5], and server systems [6]. Stroud et al. [7] reported the Malicious and Accidental Fault Tolerance for Internet Applications (MAFTIA) project, which was a three-year European research project that explored the techniques to build actual intrusion tolerant systems. The above implementation approaches are based on the redundant design at the architecture level on secure software systems. Djemaiel et al. [8] developed a cooperative intrusion detection and tolerance system (CIDTS), which took advantage of information available at multiple levels including network, host operating system (OS), and storage in early operational phases. Haridasan and Renesse [9] proposed a peer-to-peer (P2P) live-streaming system protocol built to tolerate malicious behavior at the end user level, called *SecureStream*, and showed that the proposed system could tolerate a variety

of intrusions, gracefully degrading even in the presence of a large percentage of malicious peers. For an excellent survey on this research topic, see Deswarte and Powell [10].

Because these methods can be categorized by a design diversity technique in secure systems, and are expensive to develop, the effect on implementation has to be evaluated carefully and quantitatively. To assess quantitatively security effects of computer-based systems, reliability and performance evaluation with stochastic modeling is quite effective. Littlewood et al. [11] applied fundamental techniques in reliability theory to assess the security of operational software systems, and proposed some quantitative security measures. Jonsson and Olovsson [12] also developed a quantitative method to study attacker's behavior with empirical data observed in experiments. Ortalo et al. [13] used both privilege graphs and Markov chains to evaluate system vulnerability, and derived the mean effort to security failure. Uemura and Dohi [14] focused on typical DoS attacks for a server system, and formulated the optimal patch management policy via continuous-time semi-Markov models (CTSMMs). The same authors [15] considered a secure design of an intrusion tolerant database system [16] with a control parameter, and described its stochastic behavior using a CTSMM. Recently, Zheng et al. evaluated the virtual machine based intrusion tolerant system with stochastic models [17, 18].

In this chapter, we introduce the quantitative security evaluation for the intrusion tolerance, called Scalable Intrusion Tolerant Architecture (SITAR). SITAR was developed at MCNC Inc. and at Duke University [19]. Madan et al. [20, 21] considered the security evaluation of SITAR, and described its stochastic behavior using a CTSMM. More precisely, they derived analytically the mean time to security failure (MTTSF), and investigated effects of the intrusion tolerant architecture under some attack patterns such as DoS attacks. Especially, this chapter focuses on the software maintenance with patch applying from the security point of view. Uemura et al. [22] presented the similar but somewhat different models from Madan et al. [20, 21] by introducing preventive patch management time (PPMT) which is a trigger to execute the patch applying. In commercial-off-the-shelf (COTS) distributed servers like SITAR, it is effective to patch the vulnerable server system before receiving a malicious attack. Then, in [22], the problem is to find the optimal PPMT which maximizes the steady-state system availability. Under the preventive patch management mechanism, we describe the temporal behavior of the underlying SITAR, and develop two availability models in continuous-time and discrete-time scales. On the other hand, in [23], they extended the model for SITAR to Markov regenerative Petri nets (MRSPNs) and discussed the transient analysis for quantitative security measure. This chapter mainly introduce the stochastic models for SITAR presented by Uemura et al. [22] and Fujimoto et al. [23].

The chapter is organized as follows. In Section 2, we explain SITAR and describe the stochastic behavior [20, 21]. In Section 3, we present the semi-Markov modeling to evaluate the quantitative security measure for SITAR with the embedded Markov chain (EMC) approach, and obtain the representation of respective embedded discrete-time Markov chains (DTMCs) in CTSMMs and discrete-time semi-Markov models (DTSMMs). Moreover, we derive the steady-state probability in semi-Markov processes (SMPs) (CTSMM and DTSMM) by using the mean sojourn time and the

steady-state probability in an embedded DTMC for the SMPs. In addition, we formulate the maximization problems of steady-state system availability in continuous-time and discrete-time cases, respectively, and derive analytically the optimal PPMTs maximizing the system availability. Furthermore, we derive the MTTSF for respective models by means of the EMC approach. In [Section 4](#), we introduce MRSPN based models for evaluating security. This modeling approach enables us to apply the transient analysis with the analysis tool for MRSPN. Numerical examples are presented in [Section 5](#). First we present the experiment for the optimal preventive patch management policies. It is illustrated that the preventive patch management policies can improve effectively the system availability, as well as the lifetime length of the system in some cases; and that the implementation of both preventive maintenance and intrusion tolerance may lead to keeping the system in a more reliable state. Secondly, we exhibit the transient analysis of pointwise availability under the system with and without patch management. Finally, the chapter is concluded with some remarks in [Section 6](#).

2 SITAR system

The SITAR is a COTS distributed server with an intrusion tolerant function [\[19\]](#), and consists of five major components: proxy server, acceptance monitor, ballot monitor, adaptive reconfiguration module, and audit control module. Because the usual COTS server is vulnerable for an intrusion from outside, an additional intrusion tolerant structure is introduced in SITAR. In [Fig. 1](#), it can be checked that the part denoted by a dotted square can function as an intrusion tolerance to the vulnerable end servers S_1, S_2, \dots, S_i , where P_i, B_i , and A_i in the functional blocks are the logical functions to be executed to satisfy a given service request.

Proxy servers represent public access points for the intrusion tolerant services being provided (e.g., a decision support system for military command and control, or a transaction processing system for an E-commerce site). All requests come into one of the proxy servers depending on the service needs. The proxy server enforces the service policy specified by the current intrusion tolerant strategy. The policy determines which COTS servers the request should be *forwarded* to, and how the results from these servers should be adjudicated to arrive at the final response. A new request by the proxy server to the appropriate COTS servers is made on behalf of the original client, as depicted by the thin lines from the proxy servers to the COTS servers. Relevant ballot monitors and acceptance monitors are also informed of this decision.

When the responses (signified by the thick lines from right to left in [Fig. 1](#)) are generated by the COTS servers, they are first processed by the acceptance monitors. The acceptance monitors apply certain validity check to the responses, forwarding them to the ballot monitors along with an indication of the check result. The acceptance monitors also detect signs of compromise on the COTS servers, and produce intrusion triggers for the adaptive reconfiguration module. The ballot monitors serve as “representatives” for the respective COTS servers, and decide on a final response through either a simple majority voting, or Byzantine agreement process. The actual process taken will depend on the current level of detected security threat. The final

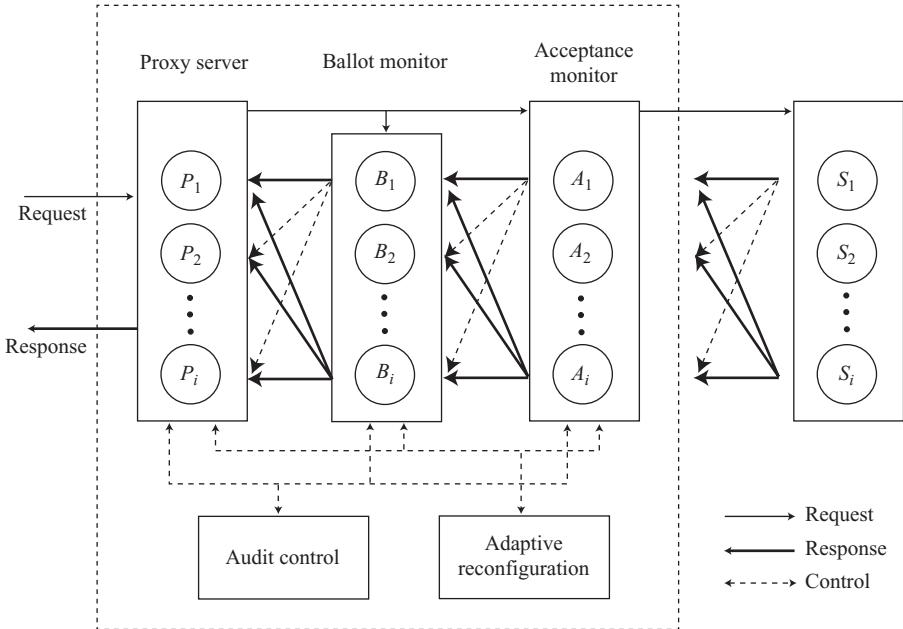


Fig. 1 Configuration of SITAR.

response is forwarded to the proxy servers to be delivered to the remote client. The adaptive reconfiguration module receives intrusion trigger information from all other modules, then evaluates intrusion threats, the tolerance objectives, and the cost/performance impact; and finally generates new configurations for the system. Because it is assumed that any individual component can be compromised, the backup adaptive reconfiguration module is provided to guard against the adaptive reconfiguration module becoming a single point of failure.

The audit control module provides means for auditing the behavior of all the other components in the intrusion tolerant system. All system modules maintain audit logs with signature protection. These logs can be verified through the audit control module. Additional diagnostic tests can be conducted through the audit control module. Intrusion triggers are distributed among three sets of modules. The triggers in the acceptance monitors are responsible for detecting compromised COTS servers. The triggers in the proxy servers are for detecting external attacks, and the triggers in the audit control will help the security administrator to monitor the secure operation of all the new functional blocks in our architecture through active auditing.

3 Semi-Markov modeling

Based on the SITAR architecture, Madan et al. [21] presented a DTSM to evaluate system security. More precisely, they provided an analytical MTTSF, and estimated effects of the system architecture on the system security quantitatively. Uemura et al. [22]

applied a security patch strategy to the SITAR architecture, and derived the system availability using steady-state analysis and the MTTSF in the similar DTSMM model to [21].

Table 1 shows the description of states of the DTSMM presented by Uemura et al. [22]. Let G be the usual state in which the COTS server can protect itself from adversaries. However, if the adversaries detect a vulnerable part, a state transition occurs from G to the vulnerable state V . Further, if adversaries attack the vulnerable part, the state moves to attack state A . Otherwise, the preventive patch is released before the attack time, and a transition to PM occurs, where a transition from PM to G is made just after the vulnerable part is completely patched. After patch management, the

Table 1 State description

State	Description
G	This state is a normal state in which COTS servers can protect themselves from adversaries. When vulnerability of COTS servers is disclosed, the system state goes to V .
V	The system is exposed to the security threat, such as compromise, exploiting vulnerability. When adversaries attack the system through the vulnerability, the state goes to A . If preventive maintenance is triggered before the attack, the system state goes to PM .
PM	The system applies security patches. After completion of the maintenance, the state becomes G .
A	The system attempts to detect an illegal action (or access). There are two possible situations. The first situation is that the illegal action cannot be detected. In this case, the state goes to UC . Otherwise, the system goes to C to diagnosis the action.
UC	The security failure occurs due to the attack by adversaries. The system is forced to undergo recovery processes (corrective maintenance). After finishing the corrective maintenance, the system state becomes G again.
C	The system begins to diagnosis the illegal action detected. At this state, the system first tries to mask the illegal access. If the masking is succeeded, the system goes to MC . Otherwise, the diagnosis still goes on at the state TR .
MC	The system provides services to users, though minor errors causing the illegal access are being fixed in the background. After completion of the repair, the system state becomes G .
TR	Several corrective inspections are tried in parallel with services. If a fatal system error is found, the system state goes to F . Otherwise, the state becomes C_2 .
F	The failure state starts a recovery operation to fix a fatal system error. After fixing, the state goes to G again.
C_2	There are two possible situations. If the damage of attack is not so large, the system state goes to GD to keep servicing to users. Otherwise, if the damage is large; for example, database recovery is needed to be repaired, the state goes to FS .
FS	The system stops servicing to users and applies recovery operation. After completion of the recovery, the state goes to G .
GD	The system keeps servicing while the quality of service is degraded. After removing system secure errors, the system state goes to G .

vulnerable part can be masked as if it was completely maintained, and the server system becomes as good as new.

On the other hand, in the attack state A , two possible states can be taken. If the problem caused by the attack cannot be resolved, and the containment of the damaged part fails, the corresponding event can be regarded as a system failure, and the initialization or reconfiguration of the system is performed as a corrective maintenance (repair) at UC . After completing that action, the system state makes a transition to G again, and becomes as good as new. Meanwhile, if the intrusion or attack is detected, then the state goes to C . In state C , one of two instantaneous transitions without time delay, which are denoted by dotted-lines in Fig. 2, can occur; i.e., if the attacked and damaged part is not so significant, and does not lead to a serious system failure directly, then the system state makes a transition from C to MC with probability $1 - p$ ($0 \leq p \leq 1$), and the damaged part can be contained by means of the fail safe function. After the containment, the system state moves back to G by masking the damaged part.

Otherwise, if the containment of the damaged part with serious effects to the system fails, the state goes to TR with probability p . We call this probability the *triage probability*. In state TR , several corrective inspections are tried in parallel with services. If the system is diagnosed as in failure, then the state moves to F , the service operation is stopped, and the recovery operation starts immediately. After completing the recovery from the system failure, the system becomes as good as new in G . Otherwise, it goes to the so-called nonfailure state denoted by C_2 . Here, two states can be taken: it may be switched to the gracefully service degradation in GD with probability q ($0 \leq q \leq 1$); or the service operation is forced to stop, and the corrective maintenance starts immediately.

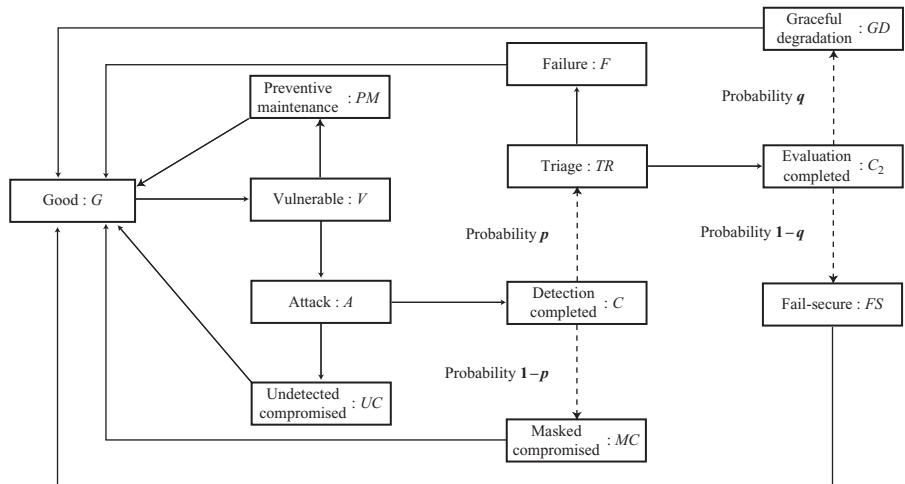


Fig. 2 Block diagram of SITAR behavior.

The main differences from Madan et al. [20, 21] are (i) the timing to trigger the preventive patch release can be turned up to improve the security effect, although Madan et al. [20, 21] did not take that into account; and (ii) in two states, C and C_2 , instantaneous transitions are allowed in the present model, although they were represented as random transitions with time delay in [20, 21].

To derive the system availability, we define the time interval from G to G as one cycle, and suppose that the same cycle repeats over an infinite time horizon. For respective states, let $F_{i,j}(t)$ and $F_{i,j}(n)$ ($i, j \in \{G, V, A, PM, UC, C, MC, TR, C_2, FS, GD, F\}$) denote the continuous and discrete transition probability distributions with p.d.f. $f_{i,j}(t)$ and p.m.f. $f_{i,j}(n)$, respectively, in CTSMM and DTSMM, where $f_{i,j}(0) = 0$ and mean $\mu_{i,j}(> 0)$. It is assumed that CTSMM and DTSMM are both irreducible and ergodic.

In Figs. 3 and 4, we give the transition diagrams of CTSMM and DTSMM, respectively. It is assumed that the preventive patch management is triggered just after t_0 (≥ 0) or n_0 ($= 0, 1, \dots$) time units elapse in the vulnerable state V in CTSMM or DTSMM. In this case, the transition probability from V to PM is given by $F_{V,PM}(t) = 1$ ($t \geq t_0$) ($F_{V,PM}(n) = 1(n \geq n_0)$) or $F_{V,PM}(t) = 0$ ($t < t_0$) ($F_{V,PM}(n) = 0(n < n_0)$). From the preliminary work above, we formulate the steady-state system availability as a function of the PPMT t_0 or n_0 .

The embedded DTMC representation of both CTSMM and DTSMM is illustrated in Fig. 5. By using the mean sojourn time at state $k \in \{G, V, PM, A, UC, MC, TR, FS,$

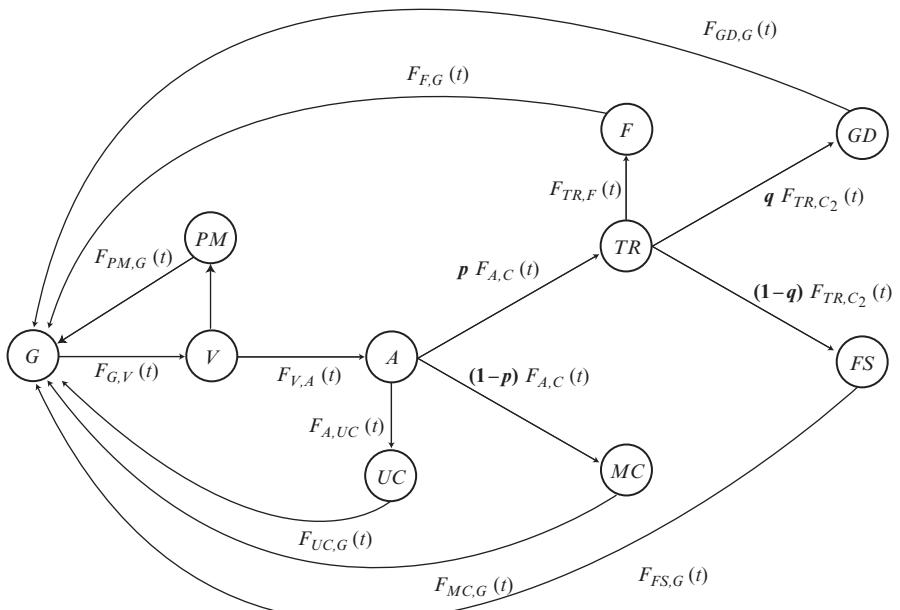


Fig. 3 Transition diagram of CTSMM.

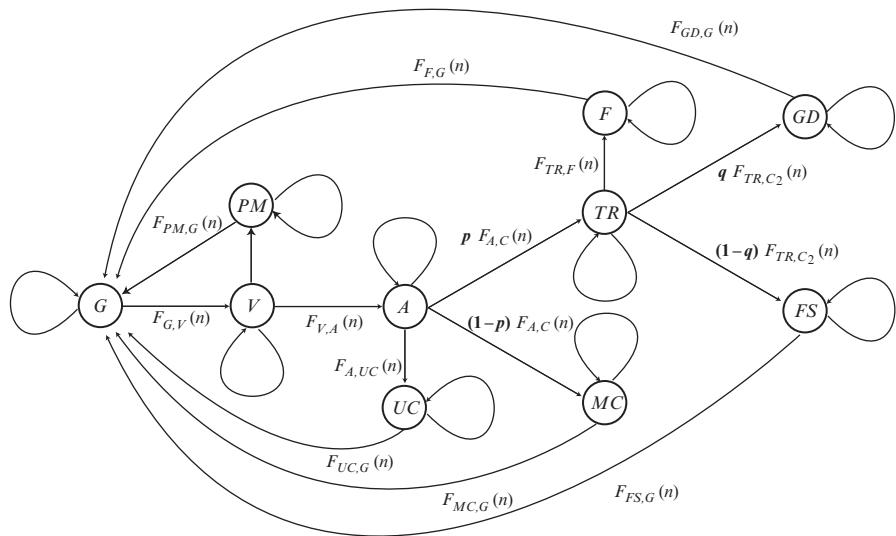


Fig. 4 Transition diagram of DTSMM.

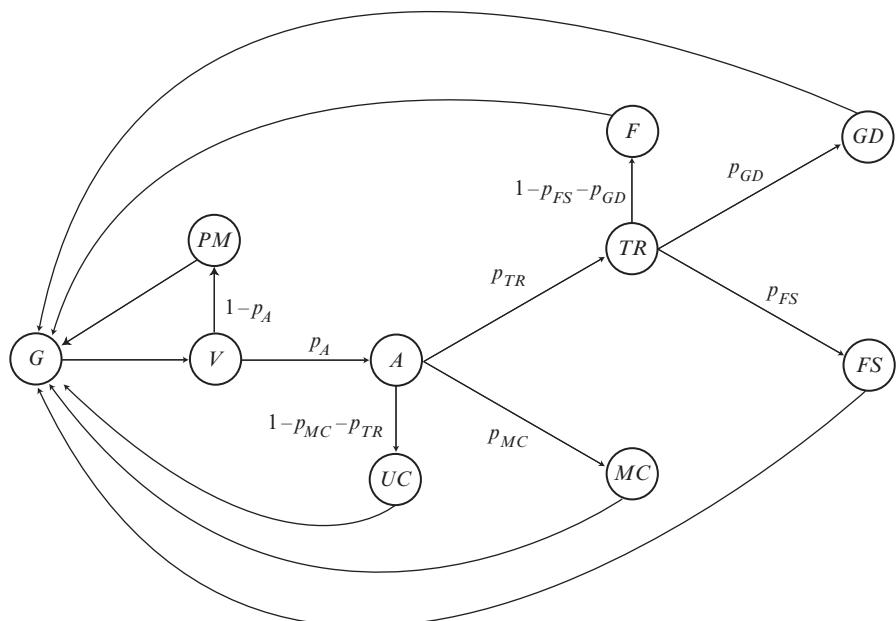


Fig. 5 EMC representation.

$GD, F\}$, h_k , and the steady-state probability in the embedded DTMC in Fig. 5, π_k , we can derive the steady-state probability π_k in SMP (CTSMM and DTSMM) in the following set of equations.

$$\pi_G = h_G/\phi, \quad (1)$$

$$\pi_V = h_V \pi_G / h_G, \quad (2)$$

$$\pi_{PM} = (1 - p_A) h_{PM} \pi_G / h_G, \quad (3)$$

$$\pi_A = p_A h_A \pi_G / h_G, \quad (4)$$

$$\pi_{UC} = p_A (1 - p_{MC} - p_{TR}) h_{UC} \pi_G / h_G, \quad (5)$$

$$\pi_{MC} = p_A p_{MC} h_{MC} \pi_G / h_G, \quad (6)$$

$$\pi_{TR} = p_A p_{TR} h_{TR} \pi_G / h_G, \quad (7)$$

$$\pi_{FS} = p_A p_{TR} p_{FS} h_{FS} \pi_G / h_G, \quad (8)$$

$$\pi_{GD} = p_A p_{TR} p_{GD} h_{GD} \pi_G / h_G, \quad (9)$$

$$\pi_F = p_A p_{TR} (1 - p_{FS} - p_{GD}) h_F \pi_G / h_G, \quad (10)$$

where

$$\begin{aligned} \phi &= h_G + h_V + (1 - p_A) h_{PM} + p_A [h_A + (1 - p_{MC} - p_{TR}) h_{UC} \\ &\quad + p_{MC} h_{MC} + p_{TR} \{h_{TR} + p_{FS} h_{FS} \\ &\quad + p_{GD} h_{GD} + (1 - p_{FS} - p_{GD}) h_F\}]. \end{aligned} \quad (11)$$

3.1 Availability formulation and optimal preventive patch management policy in continuous-time model

From Fig. 5, it is seen immediately to obtain

$$p_A = p_A(t_0) = F_{V,A}(t_0), \quad (12)$$

$$p_{MC} = (1 - p) \int_0^\infty \bar{F}_{A,UC}(t) dF_{A,C}(t), \quad (13)$$

$$p_{TR} = p \int_0^\infty \bar{F}_{A,UC}(t) dF_{A,C}(t), \quad (14)$$

$$p_{FS} = (1 - q) \int_0^\infty \bar{F}_{TR,F(t)} dF_{TR,C_2}(t), \quad (15)$$

$$p_{GD} = q \int_0^{\infty} \bar{F}_{TR,F}(t) dF_{TR,C_2}(t) \quad (16)$$

and

$$h_G = \mu_{G,V}, \quad (17)$$

$$h_V = h_V(t_0) = \int_0^{t_0} \bar{F}_{V,A}(t) dt, \quad (18)$$

$$h_{PM} = \mu_{PM,G}, \quad (19)$$

$$h_A = \int_0^{\infty} t \bar{F}_{A,UC}(t) dF_{A,C}(t) + \int_0^{\infty} t \bar{F}_{A,C}(t) dF_{A,UC}(t), \quad (20)$$

$$h_{UC} = \mu_{UC,G}, \quad (21)$$

$$h_{MC} = \mu_{MC,G}, \quad (22)$$

$$h_{TR} = \int_0^{\infty} t \bar{F}_{TR,F}(t) dF_{TR,C_2}(t) + \int_0^{\infty} t \bar{F}_{TR,C_2}(t) dF_{TR,F}(t), \quad (23)$$

$$h_{FS} = \mu_{FS,G}, \quad (24)$$

$$h_{GD} = \mu_{GD,G}, \quad (25)$$

$$h_F = \mu_{F,G}. \quad (26)$$

The steady-state system availability is defined as a fraction of time when the service by a server system can be provided continuously. Hence, the formulation of the steady-state system availability is reduced to the derivation of the mean sojourn time at each state. Note that the system down states correspond to states PM , CUC , CF , and FS , so that the steady-state system availability is represented as a function of t_0 by

$$\begin{aligned} AV(t_0) &= \pi_G + \pi_V + \pi_A + \pi_{MC} + \pi_{TR} + \pi_{GD} \\ &= U(t_0)/T(t_0), \end{aligned} \quad (27)$$

where

$$\begin{aligned} U(t_0) &= h_G + h_V(t_0) + p_A(t_0)\{h_A + p_{MC}h_{MC} + p_{TR}(h_{TR} + p_{GD}h_{GD})\} \\ &= \mu_{G,V} + \int_0^{t_0} \bar{F}_{V,A}(t) dt + \alpha F_{V,A}(t_0), \end{aligned} \quad (28)$$

$$\begin{aligned} T(t_0) &= U(t_0) + (1 - p_A(t_0))h_{PM} + p_A(t_0)[(1 - p_{MC} - p_{TR})h_{UC} \\ &\quad + p_{TR}\{p_{FS}h_{FS} + (1 - p_{FS} - p_{GD})h_F\}] \\ &= \mu_{G,V} + \int_0^{t_0} \bar{F}_{V,A}(t) dt + \mu_{PM,G}\bar{F}_{V,A}(t_0) + \beta F_{V,A}(t_0). \end{aligned} \quad (29)$$

The problem is to seek the optimal PPMT, t_0^* , maximizing the steady-state system availability $AV(t_0)$. Taking the differentiation of $AV(t_0)$ with respect to t_0 , and setting that result equal to 0 yields the nonlinear function $l(t_0) = 0$, where

$$l(t_0) = \{1 + \alpha r_{V,A}(t_0)\}T(t_0) - U(t_0)\{1 + (\beta - \mu_{PM,G})r_{V,A}(t_0)\}, \quad (30)$$

and

$$\alpha = h_A + p_{MC}h_{MC} + p_{TR}(h_{TR} + p_{GD}h_{GD}), \quad (31)$$

$$\beta = \alpha + (1 - p_{MC} - p_{TR})h_{UC} + p_{TR}\{p_{FS}h_{FS} + (1 - p_{FS} - p_{GD})h_F\}. \quad (32)$$

In the above expressions, α and β are the mean up time for one cycle and the total mean time for traveling from state A to G , respectively.

We make two parametric assumptions:

- (A-1) $\alpha + \mu_{PM,G} < \beta$ and
- (A-2) $\alpha \mu_{PM,G} < \mu_{G,V}(\beta - \alpha - \mu_{PM,G})$.

From their definitions, it is evident that $\alpha < \beta$. The assumption (A-1) implies that the sum of the mean up time after state A , and the mean preventive maintenance time based on the patch release, is strictly smaller than the total mean time. On the other hand, the assumption (A-2) seems to be somewhat technical, but is needed to guarantee a unique optimal PPMT. These two assumptions were numerically checked and could be validated in many parametric cases. We characterize the optimal PPMT, maximizing the steady-state system availability, in the following proposition.

Proposition 1.

- (A) Suppose that $F_{V,A}(t)$ is strictly increasing hazard rate (IHR) under assumptions (A-1) and (A-2).

(i) If $l(0) > 0$ and $l(\infty) < 0$, then there exists a unique optimal PPMT $t_0^*(0 < t_0^* < \infty)$ satisfying $l(t_0^*) = 0$. The corresponding steady-state system availability $AV(t_0^*)$ is given by

$$AV(t_0^*) = \frac{1 + \alpha r_{V,A}(t_0^*)}{1 + (\beta - \mu_{PM,G})r_{V,A}(t_0^*)}. \quad (33)$$

(ii) If $l(0) \leq 0$, then the optimal PPMT is $t_0^* = 0$, and the corresponding maximum steady-state system availability is given by

$$AV(0) = \frac{\mu_{G,V}}{\mu_{G,V} + \mu_{PM,G}}. \quad (34)$$

(iii) If $l(\infty) \geq 0$, then the optimal PPMT is $t_0^* \rightarrow \infty$, and the corresponding maximum steady-state system availability is given by

$$AV(\infty) = \frac{\mu_{G,V} + \mu_{V,A} + \alpha}{\mu_{G,V} + \mu_{V,A} + \beta}. \quad (35)$$

- (B) Suppose that $F_{V,A}(t)$ is decreasing hazard rate (DHR) under the assumptions (A-1) and (A-2). If $AV(0) > AV(\infty)$, then $t_0^* = 0$; otherwise $t_0^* \rightarrow \infty$.

Proof. Further differentiation of the function $l(t_0)$ leads to

$$\frac{dl(t_0)}{dt_0} = \frac{dr_{V,A}(t_0)}{dt_0} \left\{ \alpha T(t_0) - (\beta - \mu_{PM,G}) U(t_0) \right\}. \quad (36)$$

If $F_{V,A}(t)$ is strictly IHR under **(A-1)** and **(A-2)**, the right-hand side of Eq. (36) takes a negative value.

$$\begin{aligned} & \alpha T(t_0) - (\beta - \mu_{PM,G}) U(t_0) \\ &= -(\beta - \alpha - \mu_{PM,G}) \left\{ \mu_{G,V} + \int_0^{t_0} \bar{F}_{V,A}(t) dt \right\} + \alpha \mu_{PM,G} \end{aligned} \quad (37)$$

is strictly negative under the assumptions (A-1) and (A-2). Hence, the function $l(t_0)$ is a strictly decreasing function of t_0 , and $AV(t_0)$ is strictly quasiconcave in t_0 . In this situation, if $l(0) > 0$ and $l(\infty) < 0$, then there exists a unique $t_0^*(0 < t_0^* < \infty)$ maximizing $AV(t_0)$, which satisfies the nonlinear equation $l(t_0^*) = 0$. If $l(0) \leq 0$ or $l(\infty) \geq 0$, then the function $AV(t_0)$ decreases or increases, and the resulting optimal PPMT becomes $t_0^* = 0$ or $t_0^* \rightarrow \infty$. On the other hand, if $F_{V,A}(t)$ is DHR, the function $AV(t_0)$ is a quasiconvex function of t_0 , and the optimal PPMT trivially becomes $t_0^* = 0$ or $t_0^* \rightarrow \infty$. \square

3.2 Availability formulation and optimal preventive patch management policy in discrete-time model

In a fashion similar to the continuous-time model, we obtain

$$p_A = p_A(n_0) = F_{V,A}(n_0 - 1), \quad (38)$$

$$p_{MC} = (1-p) \sum_{x=0}^{\infty} \sum_{w=x}^{\infty} f_{A,UC}(w) f_{A,C}(x), \quad (39)$$

$$p_{TR} = p \sum_{x=0}^{\infty} \sum_{w=x}^{\infty} f_{A,UC}(w) f_{A,C}(x), \quad (40)$$

$$p_{FS} = (1-q) \sum_{z=0}^{\infty} \sum_{y=z}^{\infty} f_{TR,F}(y) f_{TR,C_2}(z), \quad (41)$$

$$p_{GD} = q \sum_{z=0}^{\infty} \sum_{y=z}^{\infty} f_{TR,F}(y) f_{TR,C_2}(z) \quad (42)$$

and

$$h_G = \mu_{G,V}, \quad (43)$$

$$h_V = h_V(n_0) = \sum_{n=0}^{n_0-1} \bar{F}_{V,A}(n), \quad (44)$$

$$h_{PM} = \mu_{PM,G}, \quad (45)$$

$$h_A = \sum_{x=0}^{\infty} \sum_{w=0}^{x-1} wf_{A,UC}(w)f_{A,C}(x) + \sum_{x=0}^{\infty} \sum_{w=x}^{\infty} xf_{A,UC}(w)f_{A,C}(x), \quad (46)$$

$$h_{UC} = \mu_{UC,G}, \quad (47)$$

$$h_{MC} = \mu_{MC,G}, \quad (48)$$

$$h_{TR} = \sum_{z=0}^{\infty} \sum_{y=0}^{z-1} yf_{TR,F}(y)f_{TR,C_2}(z) + \sum_{z=0}^{\infty} \sum_{y=z}^{\infty} zf_{TR,F}(y)f_{TR,C_2}(z), \quad (49)$$

$$h_{FS} = \mu_{FS,G}, \quad (50)$$

$$h_{GD} = \mu_{GD,G}, \quad (51)$$

$$h_F = \mu_{F,G}. \quad (52)$$

Define the steady-state system availability as a function of n_0 :

$$\begin{aligned} AV(n_0) &= \pi_G + \pi_V + \pi_A + \pi_{MC} + \pi_{TR} + \pi_{GD} \\ &= U(n_0)/T(n_0), \end{aligned} \quad (53)$$

where

$$U(n_0) = \mu_{G,V} + \sum_{n=0}^{n_0-1} \bar{F}_{V,A}(n) + \alpha F_{V,A}(n_0 - 1), \quad (54)$$

$$T(n_0) = \mu_{G,V} + \sum_{n=0}^{n_0-1} \bar{F}_{V,A}(n) + \mu_{PM,G} \bar{F}_{V,A}(n_0 - 1) + \beta F_{V,A}(n_0 - 1). \quad (55)$$

Taking the difference of $AV(n_0)$ with respect to n_0 , we define

$$\begin{aligned} l(n_0) &= \frac{AV(n_0 + 1) - AV(n_0)}{T(n_0 + 1)T(n_0)\bar{F}_{V,A}(n_0 - 1)} \\ &= \{1 + (\alpha - 1)r_{V,A}(n_0)\}T(n_0) - U(n_0)\{1 + (\beta - \mu_{PM,G} - 1)r_{V,A}(n_0)\}. \end{aligned} \quad (56)$$

Under the two parametric assumptions in [Section 3.1](#), we characterize the optimal PPMT in a discrete-time model.

Proposition 2.

(A) Suppose that $F_{V,A}(n)$ is strictly IHR under (A-1) and (A-2).

- (i) If $l(0) > 0$ and $l(\infty) < 0$, then there exists (at least one, at most two) optimal PPMT $n_0^*(0 < n_0^* < \infty)$ satisfying the simultaneous inequalities $l(n_0^* - 1) > 0$ and $l(n_0^*) \leq 0$. The corresponding steady-state system availability $AV(n_0^*)$ must satisfy

$$K(n_0^* + 1) \leq AV(n_0^*) < K(n_0^*), \quad (57)$$

where

$$K(n) = \frac{1 + (\alpha - 1)r_{V,A}(n)}{1 + (\beta - \mu_{PM,G} - 1)r_{V,A}(n)}. \quad (58)$$

- (ii) If $l(0) \leq 0$, then the optimal PPMT is $n_0^* = 0$, and the corresponding maximum steady-state system availability is given by

$$AV(0) = \frac{\mu_{G,V}}{\mu_{G,V} + \mu_{PM,G}}. \quad (59)$$

- (iii) If $l(\infty) \geq 0$, then the optimal PPMT is $n_0^* \rightarrow \infty$, and the corresponding maximum steady-state system availability is given by

$$AV(0) = \frac{\mu_{G,V} + \mu_{V,A} + \alpha}{\mu_{G,V} + \mu_{V,A} + \beta}. \quad (60)$$

(B) Suppose that $F_{V,A}(n)$ is DHR under (A-1) and (A-2). If $AV(0) > AV(\infty)$, then $n_0^* = 0$. Otherwise, $n_0^* \rightarrow \infty$.

Proof. Taking the difference of Eq. (56), we obtain

$$\begin{aligned} l(n_0 + 1) - l(n_0) &= \{r_{V,A}(n_0 + 1) - r_{V,A}(n_0)\} \\ &\times \left[(\alpha - \beta + \mu_{PM,G}) \left\{ \mu_{G,V} + \sum_{n=0}^{n_0-1} \bar{F}_{V,A}(n) \right\} \right. \\ &\quad \left. + \alpha \mu_{PM,G} + (\alpha - \beta) \right]. \end{aligned} \quad (61)$$

If $F_{V,A}(n)$ is strictly IHR, the right-hand side of Eq. (61) is strictly negative under (A-1) and (A-2); and the function $l(n_0)$ is strictly decreasing in n_0 . Because the steady-state system availability $AV(n_0)$ is strictly quasiconcave in n_0 in the discrete sense, if $l(0) > 0$ and $l(\infty) < 0$, then there exists at least one and at most two, optimal PPMT $n_0^*(0 < n_0^* < \infty)$ so as to satisfy $l(n_0^* - 1) > 0$ and $l(n_0^*) \leq 0$, which lead to the inequalities in Eq. (57). If $l(0) \leq 0$ or $l(\infty) \geq 0$, then the function $AV(n_0)$ decreases or increases, and the resulting optimal PPMT becomes $n_0^* = 0$ or $n_0^* \rightarrow \infty$. On the other hand, if $F_{V,A}(n)$ is DHR, then the function $AV(n_0)$ is a quasiconvex function of n_0 in the discrete sense, and the optimal PPMT is given by $n_0^* = 0$ or $n_0^* \rightarrow \infty$. \square

3.3 MTTSF

Next, we derive MTTSF [21]. Let X_C and X_Q denote the absorbing states and the transient states, respectively, for both CTSMM and DTSMM. Let

$$\mathbf{P} = \begin{bmatrix} \mathbf{Q} & \mathbf{C} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \quad (62)$$

be the whole transition probability matrix, where \mathbf{Q} and \mathbf{C} denote the transient and the absorbing probability matrices for $X_Q = \{UC, FS, GD, F\}$ and $X_C = \{G, V, PM, A, MC, TR\}$ in Fig. 5:

$$\mathbf{Q} = \begin{bmatrix} G & V & PM & A & MC & TR \\ G & 0 & 1 & 0 & 0 & 0 \\ V & 0 & 0 & \bar{p}_A & p_A & 0 \\ PM & 1 & 0 & 0 & 0 & 0 \\ A & 0 & 0 & 0 & 0 & p_{MC} \\ MC & 1 & 0 & 0 & 0 & 0 \\ TR & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (63)$$

and

$$\mathbf{C} = \begin{bmatrix} UC & FS & GD & F \\ G & 0 & 0 & 0 \\ V & 0 & 0 & 0 \\ PM & 0 & 0 & 0 \\ A & \frac{p_{MC} + p_{TR}}{p_{MC} + p_{TR}} & 0 & 0 \\ MC & 0 & 0 & 0 \\ TR & 0 & p_{FS} & p_{GD} & \frac{p_{MC} + p_{TR}}{p_{MC} + p_{TR}} \end{bmatrix} \quad (64)$$

where \mathbf{O} and \mathbf{I} are the zero matrix whose elements are 0 and the identity matrix, respectively, $\bar{p}_A = 1 - p_A$, $\overline{p_{MC} + p_{TR}} = 1 - p_{MC} - p_{TR}$, $\overline{p_{FS} + p_{GD}} = 1 - p_{FS} - p_{GD}$. Using the mean visit number V_i , and the mean sojourn time h_i in state i , MTTSF is defined by

$$\text{MTTSF} = \sum_{i \in X_t} V_i h_i, \quad (65)$$

where V_i is the solution of the simultaneous equations

$$V_i = q_i + \sum_j V_j q_{ji}, \quad i, j \in X_t, \quad (66)$$

and q_{ji} denote the elements of \mathbf{Q} . For the initial probability vector in Eq. (66), we set

$$\mathbf{q}_0 = [q_i] = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]. \quad (67)$$

Finally, solving Eq. (66) yields the mean visit number

$$V_G = \frac{1}{p_A(t_0)(1-p_{MC})}, \quad (68)$$

$$V_V = V_G, \quad (69)$$

$$V_{PM} = \bar{p}_A(t_0)V_G, \quad (70)$$

$$V_A = p_A(t_0)V_G, \quad (71)$$

$$V_{MC} = p_A(t_0)p_{MC}V_G, \quad (72)$$

$$V_{TR} = p_A(t_0)p_{TR}V_G \quad (73)$$

and leads to the analytical derivation of MTTSF.

4 MRSPN modeling

4.1 MRSPN

Petri net (PN) is a directed bipartite graph with two types of nodes: place and transitions. Places and transitions in PNs are represented by circles and rectangles, respectively. Directed arcs connect places to transitions and transitions to places. The place that connects to a transition is called an input place of the transition. On the other hand, the place that is connected from a transition is called an output place of the transition. Tokens (represented by dots) are located at places in PNs. When a transition fires, it removes a token from each input place of the transition, and puts a token to each output place of the transition. The firing of a transition occurs only when there is at least one token for each input place of the transition. Then the transition is said to be enable. A marking of a PN is given by a vector that represents the number of tokens for all the places. In the PN modeling, markings provide the state of a target system.

A stochastic PN (SPN) is defined as the PN which has random firing times. In a deterministic PN, transitions immediately fire when the transitions are enable. SPNs allow random delay times of firing from the time when transitions are enable. In particular, random firing times are allowed to obey exponential distributions. The transition with such a random firing time is called an EXP transition. On the other hand, Markov regenerative stochastic Petri nets (MRSPNs) are a versatile tool for model-based performance evaluation more than SPNs [24]. A MRSPN is defined as a superclass of SPNs, and it has at most one GEN transition, which has the random firing time following a general probability distribution.

The analysis of MRSPN is based on the analysis of Markov regenerative process (MRGP). The MRGP is one of the widest classes of stochastic point process, which has two time points; regenerative and nonregenerative points. Roughly speaking, the regenerative point renews the general probability distribution for a state

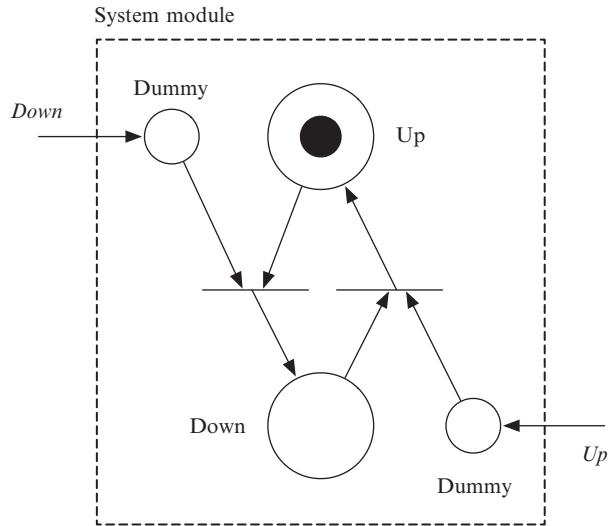
transition and the nonregenerative point does not affect the age of the general probability distribution. The MRGPs exactly includes the semi-Markov model where the state transition is triggered by general distributions. Thus the MRSPN modeling can also handle the semi-Markov modeling. The advantage of MRSPN modeling is to utilize the computer-aided tools to analyze the MRSPN. Especially, by analyzing the structure of MRSPN and utilizing some mathematical technique such as phase expansion, we can apply the transient analysis of system. This gives us the different aspect of security evaluation from the stationary analysis using the semi-Markov models.

4.2 Description of MRSPNs for the SITAR

To build an MRSPN-based model, we consider the modularized approach for the system behavior. Bobbio et al. [25] presented a unified modeling framework using a fluid stochastic Petri net for a general dependability system with checkpointing, rejuvenation, restart, and reboot. In [25], the system is divided into some major functional modules. At each module, they provide components described by a fluid stochastic Petri net. Similar to [25], we also propose a general modeling framework of an intrusion tolerant system with major analysis modules, and purposes MRSPN-based models for the major modules.

Concretely, we consider the following four MRSPN modules to model an intrusion tolerant system:

System module	This analysis module represents system behavior related to security measures. More precisely, the module has two states: up and down. The module has two input arcs to switch the state from up to down and vice versa. The security measure such as MTTSF can be measured by observing the state of this module.
Vulnerability module	The vulnerability module provides the modeling of security threat (attack and/or compromise) occurrence. The module has an input arc which is a trigger of removing vulnerability (vulnerability-free), and an output arc which represents the security threat occurrence. For example, removing the vulnerability is caused by releasing and applying a security patch. That is, this module governs a random firing time to a security threat occurrence.
Intrusion tolerance module	The intrusion tolerance module determines the system operation after a security threat occurs. There are one input arc and two output arcs. The input arc receives a signal for a threat occurrence event, i.e., a trigger of starting an intrusion tolerant process. An output arc sends a signal to change the system state to down. Another is a notification of recovery process completion.
Maintenance module	The maintenance module is similar to the intrusion tolerance module. The maintenance module also has one input arc and two output arcs. The three arcs have the same functions as those in the intrusion tolerance module. The intrusion tolerance and maintenance modules essentially correspond to preventive and corrective actions for security threats. Thus the maintenance module, for example, represents a security patching action of the system.

Fig. 6 System module.

According to Madan et al. [21] and Uemura et al. [22], we design the above four major modules. Fig. 6 presents an MRSPN of System module. System module has four places, two immediate transitions, and one token. Immediate transitions fire just after tokens are put into all the input places. Therefore, when the module receives a *down* signal, the token immediately moves to the down place. Figs. 7 and 8 present MRSPNs of Vulnerability modules for Madan et al. [21] and Uemura et al. [22], respectively. The module has four places, four or five transitions, and one token. The token represents three vulnerability levels: vulnerability-free, vulnerable, and threat occurrence. The transitions given by rectangles are EXP transitions which fire with random times exponentially distributed. When the transition located between the vulnerable and threat occurrence places fires, a signal *threat occurrence* is sent to external modules. When a signal *vulnerability-free* is received, the token moves to the vulnerability-free place. The main difference between Madan et al.'s and Uemura et al.'s models is whether the transition from the vulnerable to the vulnerability-free exists or not. Fig. 9 depicts an MRSPN of Intrusion tolerance module. In this figure, GEN transitions with the general distributed firing times (represented by bars) are used. Moreover, the two places, "detection completion" and "evaluation completion" have competitive immediate transitions. When two or more immediate transitions are competitive, a firing transition is randomly selected based on the weight of each transition. This module represents the status of progress of an intrusion tolerant process. When the token from an external module reaches to the places, detection failure, failure, and fail-secure, a *system down* signal is sent to external modules. When each of recovery processes: detection failure, masking, failure, fail-secure, and graceful degradation, is completed, a signal *recovery completion* is sent to external modules. Finally, Fig. 10 presents an MRSPN of Maintenance module. The module has two places and two GEN transitions. The first GEN transition corresponds to the time interval of maintenance, the second one the total time for maintenance.

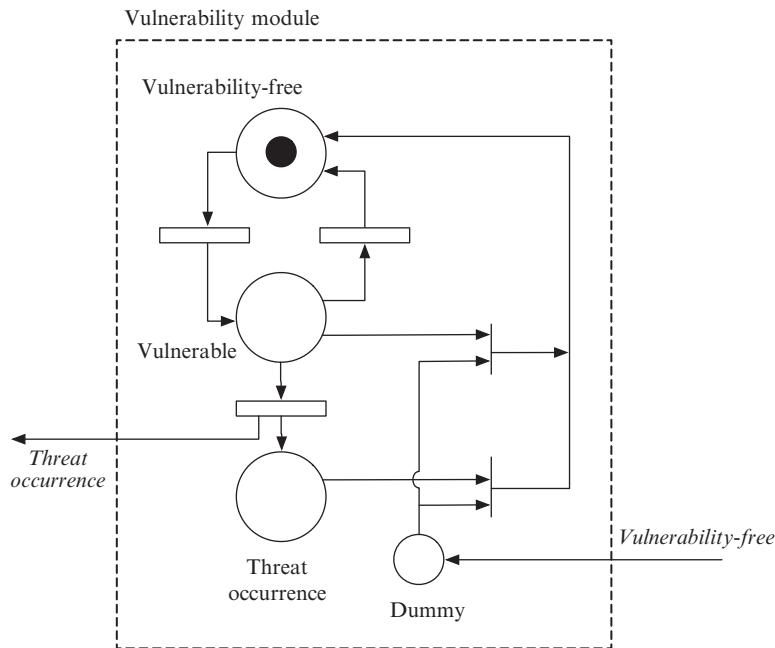


Fig. 7 Vulnerability module (type 1).

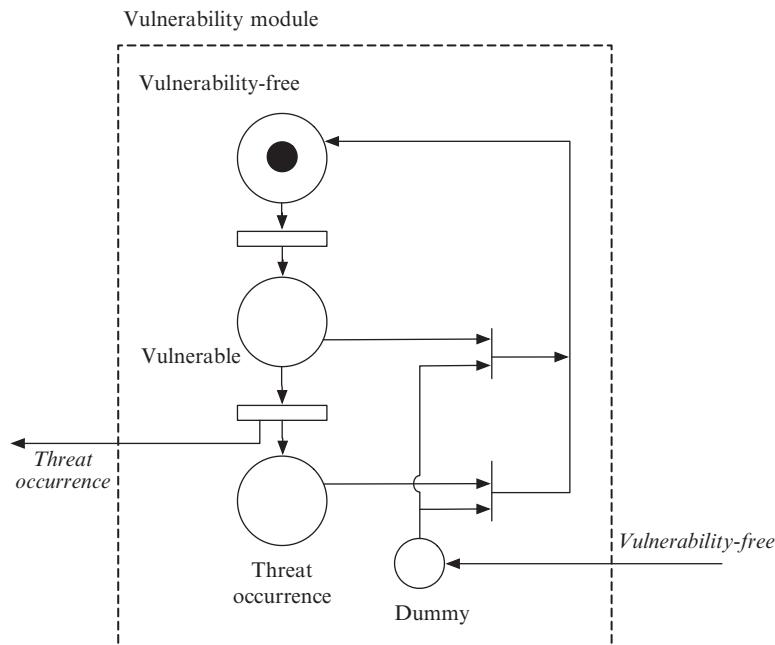


Fig. 8 Vulnerability module (type 2).

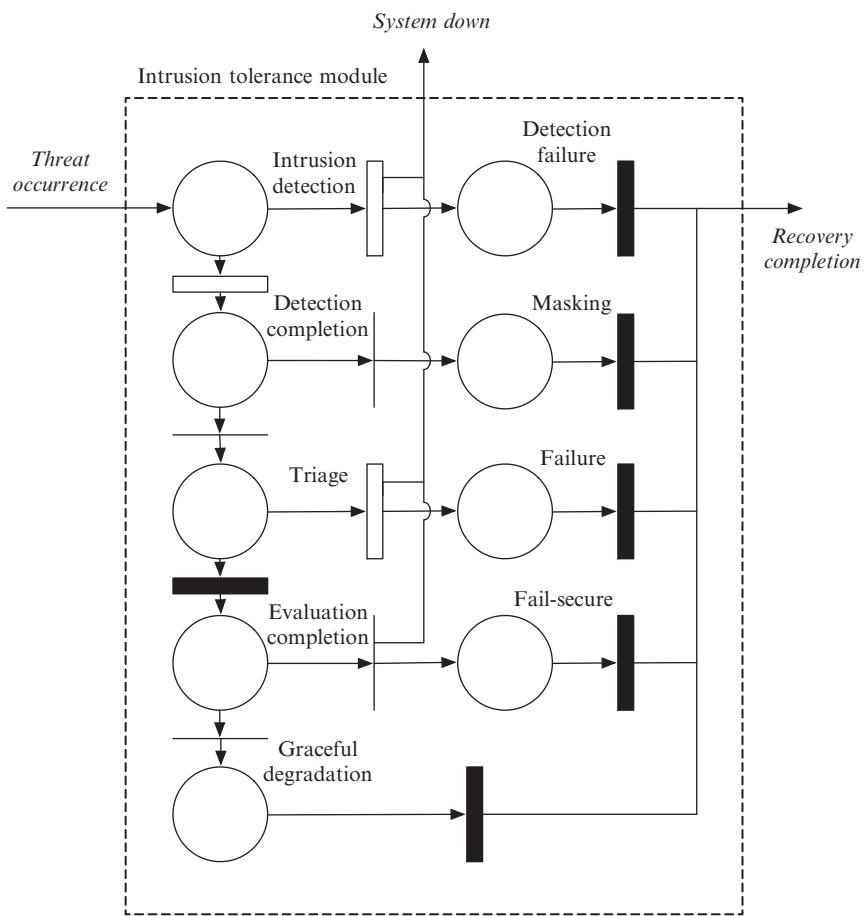
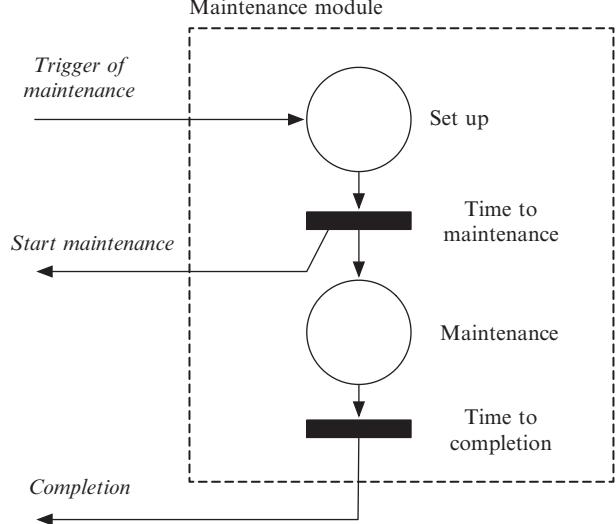


Fig. 9 Intrusion tolerance module.

Fig. 10 Maintenance module.



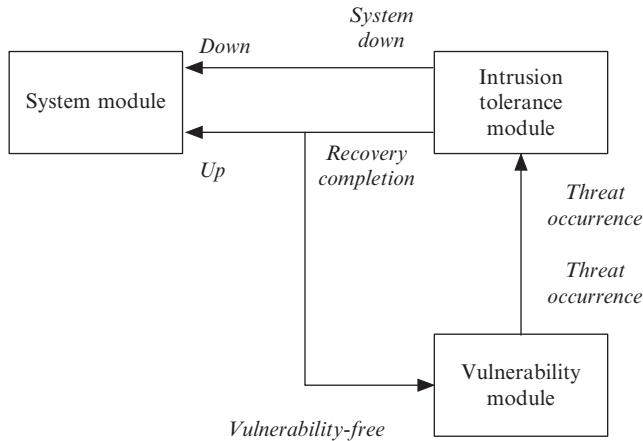


Fig. 11 MRSPN-based Madan et al.’s SITAR-security model.

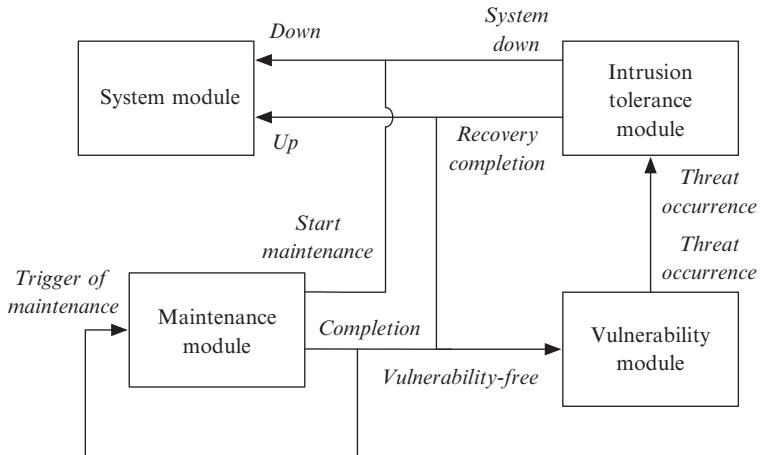


Fig. 12 MRSPN-based Uemura et al.’s SITAR-security model.

By combining the above four modules, the security evaluation models by Madan et al. [21] and Uemura et al. [22] are given in Figs. 11 and 12. Although these MRSPN-based models essentially indicate the same stochastic behavior as original DTSMMs, these more clarify the difference between Madan et al.’s and Uemura et al.’s models than the DTSMM model presented in Section 3. Moreover, this modeling framework helps us to extend the model architecture. For example, if we consider multilevel vulnerable states, we can make an MRSPN model to replace the original module by Vulnerability module including multilevel vulnerable states. Since Madan et al.’s and Uemura et al.’s models do not allow two or more security threat

occurrences simultaneously. Our framework enables simply to give such a complex model. In addition, we can apply existing numerical techniques or tools to analyze MRSPNs directly. These features provide much of benefits to the model-based security evaluation.

5 Numerical examples

5.1 Security evaluation with semi-Markov models

In this section, we derive the optimal PPMTs (t_0^* and n_0^*) characterized in [Sections 3.1](#) and [3.2](#); and quantify two security measures: steady-state system availability and MTTSF. Suppose the following parametric circumstance. $\mu_{G,V} = 72$, $\mu_{PM,G} = 10$, $\mu_{A,UC} = 8$, $\mu_{UC,G} = 24$, $\mu_{A,C} = 12$, $\mu_{MC,G} = 12$, $\mu_{TR,F} = 6$, $\mu_{F,G} = 48$, $\mu_{TR,C_2} = 8$, $\mu_{FS,G} = 30$, and $\mu_{GD,G} = 40$. We are concerned with the following four cases.

- (i) **Case 1:** $p = 0$, i.e., the system state makes a transition from C to MC with probability one.
- (ii) **Case 2:** $p = 0.5$ and $q = 0.5$.
- (iii) **Case 3:** $p = 1$ and $q = 0$, i.e., the service operation at C_2 is forced to stop with probability one.
- (iv) **Case 4:** $p = 1$ and $q = 1$, i.e., the gracefully degradation can be observed with probability one.

5.1.1 Continuous-time model

Suppose that $f_{V,A}(t)$ is the gamma p.d.f. with shape parameter k and scale parameter d .

$$f_{V,A}(t) = t^{k-1} \frac{\exp\{-t/d\}}{\Gamma(k)d^k} dt, \quad (74)$$

where $\Gamma(\cdot)$ denotes the standard gamma function. [Table 2](#) presents the optimal PPMT, its associated system availability and MTTSF for varying r in four operation circumstances (Case 1 ~ Case 4), where $d = 9$ and “ Δ ” in the table denotes the increment from the non-PPMT case ($t_0 \rightarrow \infty$). The system availability could be improved, especially, up to 18.5% in Case 3. The main reason why this observation could be obtained was that services frequently stopped in Case 3. In [Table 3](#), we calculate the optimal PPMT, and the corresponding system availability with varying d and fixed $r = 3$. From this result, it can be seen that the preventive maintenance is quite effective. In addition, as the value of d increases, i.e., the time to an attack since the detection of vulnerability is longer, the steady-state system availability monotonically increases. The remarkable case with highest increment was also Case 3.

In [Tables 4](#) and [5](#), we derive the MTTSF with the optimal PPMT maximizing system availability. The MTTSF could be improved even by controlling the PPMT for system availability, and its effect on MTTSF was about 9% at the minimum in Case 2 ([Table 5](#)). In the maximum case, the optimal PPMT provided a MTTSF that is 392

Table 2 Dependence of steady-state system availability on parameter k in continuous-time operation

k	Case 1			Case 2		
	t_0^*	$AV(t_0^*)$	Δ	t_0^*	$AV(t_0^*)$	Δ
1	0	0.8780	1.7606	0	0.8780	8.3231
2	5.7397	0.8814	0.8878	1.2955	0.8790	6.5797
3	15.4761	0.8888	0.6610	6.0877	0.8833	5.5493
4	24.9298	0.8960	0.5765	12.2036	0.8891	4.9136
5	33.9538	0.9027	0.5373	18.8321	0.8951	4.4778
k	Case 3			Case 4		
	t_0^*	$AV(t_0^*)$	Δ	t_0^*	$AV(t_0^*)$	Δ
1	0	0.8780	18.5409	0	0.8780	11.4970
2	0.6084	0.8785	15.7226	0.9069	0.8787	9.5082
3	3.8974	0.8816	13.7366	4.9151	0.8824	8.2243
4	8.7152	0.8864	12.3337	10.3650	0.8877	7.3821
5	14.2706	0.8919	11.2806	16.4421	0.8935	6.7795

Table 3 Dependence of steady-state system availability on parameter d in continuous-time operation

k	Case 1			Case 2		
	t_0^*	$AV(t_0^*)$	Δ	t_0^*	$AV(t_0^*)$	Δ
5	5.2862	0.8824	1.3903	2.3310	0.8802	7.3083
10	18.8823	0.8906	0.5461	7.2456	0.8842	5.1938
20	72.7078	0.9084	0.0768	22.9515	0.8946	2.8367
50	∞	0.9414	0	99.8380	0.9234	0.6790
100	∞	0.9612	0	269.9410	0.9497	0.0074
k	Case 3			Case 4		
	t_0^*	$AV(t_0^*)$	Δ	t_0^*	$AV(t_0^*)$	Δ
5	1.5358	0.8795	16.7533	1.9124	0.8798	10.2982
10	4.6109	0.8822	13.1053	5.8299	0.8832	7.7946
20	13.9951	0.8894	8.5856	18.0010	0.8919	4.7948
50	58.4712	0.9128	3.2153	75.9138	0.9181	1.5499
100	156.4120	0.9390	0.5487	201.8110	0.9444	0.1483

Table 4 Dependence of MTTSF on parameter k in continuous-time operation

k	Case 1			Case 2		
	t_0^*	MTTSF	Δ	t_0^*	MTTSF	Δ
1	0	∞	∞	0	∞	∞
2	5.74	1.08E+03	552.52	1.30	1.11E+04	8933.27
3	15.48	6.46E+02	257.02	6.09	3.51E+03	2526.52
4	24.93	5.78E+02	195.15	12.20	2.40E+03	1558.57
5	33.95	5.75E+02	172.71	18.83	2.05E+03	1210.45
k	Case 3			Case 4		
	t_0^*	MTTSF	Δ	t_0^*	MTTSF	Δ
1	0	∞	∞	0	∞	∞
2	0.61	3.78E+04	39,220.60	0.91	1.75E+04	18,052.40
3	3.90	8.75E+03	8215.60	4.92	4.79E+03	4454.07
4	8.72	5.30E+03	4538.94	10.37	3.10E+03	2615.84
5	14.27	4.20E+03	3307.93	16.44	2.56E+03	1976.31

Table 5 Dependence of MTTSF on parameter d in continuous-time operation

d	Case 1			Case 2		
	t_0^*	MTTSF	Δ	t_0^*	MTTSF	Δ
5	5.29	1.59E+03	889.58	2.33	8.81E+03	7332.02
10	18.88	5.63E+02	202.44	7.25	2.99E+03	2079.62
20	72.71	3.15E+02	33.33	22.95	1.19E+03	579.72
50	∞	4.04E+02	0	99.84	6.60E+02	119.29
100	∞	9.09E+02	0	269.94	7.41E+02	8.99
d	Case 3			Case 4		
	t_0^*	MTTSF	Δ	t_0^*	MTTSF	Δ
5	1.54	2.17E+04	23,225.30	1.91	1.20E+04	12,726.40
10	4.61	7.46E+03	6792.58	5.83	4.09E+03	3677.08
20	14.00	2.81E+03	1931.05	18.00	1.58E+03	1044.73
50	58.47	1.21E+03	406.60	75.91	7.79E+02	225.76
100	156.41	1.10E+03	102.28	201.81	7.93E+02	46.27

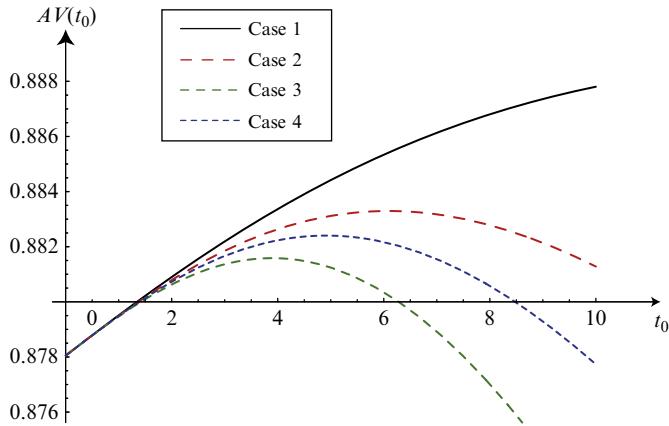


Fig. 13 Steady-state system availability in respective cases (continuous-time model).

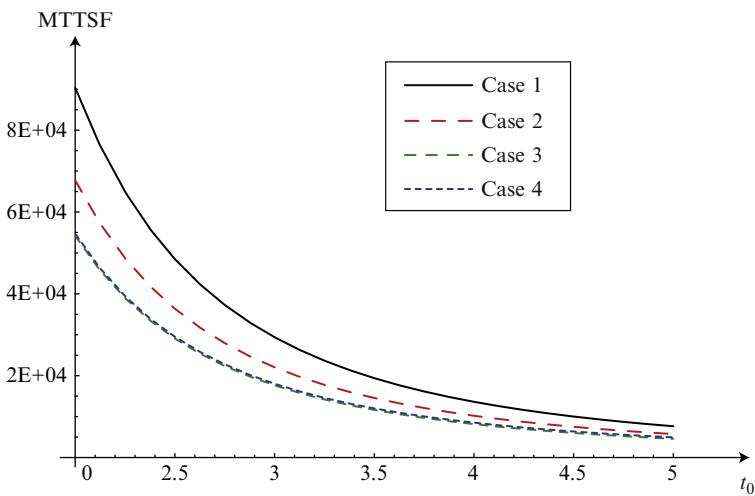


Fig. 14 MTTSF in respective cases (continuous-time model).

times longer than the optimal PPMT for that case. Figs. 13 and 14 illustrate the behavior of the system availability and MTTSF, respectively. From these figures, we know that Case 3 with many service stops gives lower system availability and MTTSF, and that its decreasing rate is remarkable. Case 4 corresponds to the well-known DoS attack with G , V , PM , UC , TR , GD , and F , where the states MC and FS can be regarded as a security failure state under the DoS attack circumstance [20, 21]. The numerical results suggest that controlling the optimal PPMT may be effective to keep the high level of system availability and to improve the system survivability.

5.1.2 Discrete-time model

Suppose that $f_{V,A}(n)$ is given by the negative binomial p.m.f.

$$f_{V,A}(n) = \binom{n-1}{r-1} \xi^r (1-\xi)^{n-r}, \quad (75)$$

where $\xi \in (0, 1)$ and $r = 1, 2, \dots$. [Table 6](#) presents the dependence of optimal PPMT and its associated system availability on the parameter r under four different scenarios. By performing the preventive patch management at a suitable timing, and comparing to the case without preventative maintenance, see that the steady-state system availability can be improved. Especially, in Case 3, the system availability could be improved up to 18.3%. Further, we execute the sensitivity analysis of optimal preventive maintenance policy on ξ in [Table 7](#). Observe that the system availability monotonically decreased as ξ increased, and that the increment of system availability was remarkable in Case 3 and Case 4.

[Tables 8](#) and [9](#) present the MTTSF with the optimal PPMT maximizing the system availability. In the discrete-time case, we could check that the MTTSF was improved to 6% in the worst case, and became 113 times longer. In [Figs. 15](#) and [16](#), we investigate the behavior of the system availability and MTTSF in discrete-time cases. Similar to the continuous-time case, Case 3 with many stops of services decreases both system availability and MTTSF. In the DoS attack case

Table 6 Dependence of steady-state system availability on parameter r in discrete-time operation

r	Case 1			Case 2		
	n_0^*	$AV(n_0^*)$	Δ	n_0^*	$AV(n_0^*)$	Δ
1	1	0.8795	1.2200	1	0.8795	7.7659
2	9	0.8834	0.4833	2	0.8800	6.0581
3	21	0.8910	0.3489	7	0.8839	5.0588
4	31	0.8983	0.3088	13	0.8895	4.4589
5	41	0.9048	0.2947	20	0.8954	4.0545
r	Case 3			Case 4		
	n_0^*	$AV(n_0^*)$	Δ	n_0^*	$AV(n_0^*)$	Δ
1	1	0.8795	18.3327	1	0.8795	10.5623
2	1	0.8795	15.5314	2	0.8796	8.6410
3	4	0.8820	13.5188	6	0.8830	7.4244
4	9	0.8865	12.1178	11	0.8881	6.6463
5	14	0.8917	11.0693	17	0.8937	6.0971

Table 7 Dependence of steady-state system availability on parameter ξ in discrete-time operation

ξ	Case 1			Case 2		
	n_0^*	$AV(n_0^*)$	Δ	n_0^*	$AV(n_0^*)$	Δ
0.01	∞	0.9619	0	278	0.9502	0.0036
0.02	∞	0.9429	0	102	0.9237	0.6167
0.05	90	0.9100	0.0243	23	0.8944	2.7007
0.1	21	0.8910	0.3489	7	0.8839	5.0588
0.2	5	0.8822	1.1139	2	0.8802	7.2196
ξ	Case 3			Case 4		
	n_0^*	$AV(n_0^*)$	Δ	n_0^*	$AV(n_0^*)$	Δ
0.01	155	0.9387	0.5690	208	0.9450	0.1199
0.02	57	0.9124	3.2971	78	0.9185	1.4069
0.05	13	0.8890	8.8217	18	0.8919	4.4908
0.1	4	0.8820	13.5188	6	0.8830	7.4244
0.2	1	0.8795	17.3377	2	0.8799	9.9234

Table 8 Dependence of MTTSF on parameter r in discrete-time operation

r	Case 1			Case 2		
	n_0^*	MTTSF	Δ	n_0^*	MTTSF	Δ
1	1	∞	∞	1	∞	∞
2	9	6.92E+02	296.12	2	1.06E+04	8351.76
3	21	4.55E+02	139.23	7	2.95E+03	2052.51
4	31	4.44E+02	115.61	13	2.16E+03	1352.50
5	41	4.47E+02	102.20	20	1.76E+03	998.77
r	Case 3			Case 4		
	n_0^*	MTTSF	Δ	n_0^*	MTTSF	Δ
1	1	∞	∞	1	∞	∞
2	1	∞	∞	2	8.40E+03	8501.41
3	4	1.00E+04	9319.42	6	3.42E+03	3107.78
4	9	4.90E+03	4141.64	11	2.72E+03	2249.19
5	14	4.33E+03	3373.56	17	2.29E+03	1734.05

Table 9 Dependence of MTTSF on parameter ξ in discrete-time operation

ξ	Case 1			Case 2		
	n_0^*	MTTSF	Δ	n_0^*	MTTSF	Δ
0.01	∞	9.35E+02	0	278	7.26E+02	6.26
0.02	∞	4.15E+02	0	102	6.39E+02	111.41
0.05	90	2.79E+02	15.48	23	1.14E+03	548.02
0.1	21	4.55E+02	139.23	7	2.95E+03	2052.51
0.2	5	1.52E+03	823.21	2	1.33E+04	11,144.40
ξ	Case 3			Case 4		
	n_0^*	MTTSF	Δ	n_0^*	MTTSF	Δ
0.01	155	1.10E+03	104.33	208	7.59E+02	41.12
0.02	57	1.24E+03	423.60	78	7.32E+02	208.80
0.05	13	3.15E+03	2208.41	18	1.49E+03	993.37
0.1	4	1.00E+04	9319.42	6	3.42E+03	3107.78
0.2	1	∞	∞	2	1.05E+04	11,357.20

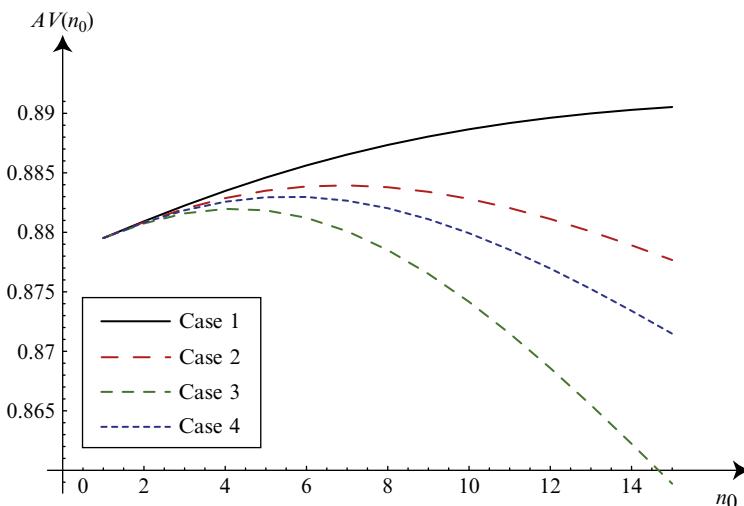


Fig. 15 Steady-state system availability in respective cases (discrete-time model).

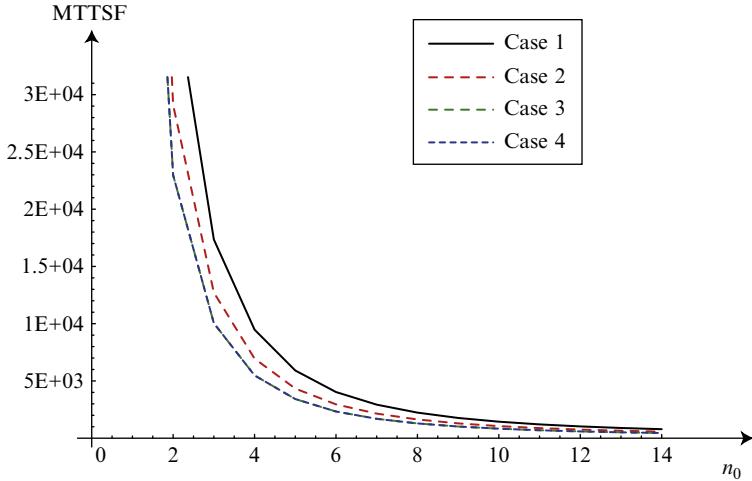


Fig. 16 MTTSF in respective cases (discrete-time model).

(Case 4), observe that controlling the optimal PPMT could keep the system availability high and the MTTSF long. Hence, implementation of intrusion detection and intrusion tolerance functions can result in robust availability and survivability in the COTS server system.

5.2 Transient analysis with MRSPN modeling

This section illustrates a numerical analysis for the MRSPN-based SITAR-security models. In particular, we apply a supplementary variable method for MRSPNs [26, 27] to compute point availability for the SITAR system as an example of transient analysis.

In this example, Vulnerability module is modeled by multilevel vulnerable states presented in Fig. 17. In this analysis module, the number of tokens corresponds to the vulnerable level, and a security threat occurs when three tokens are located at the vulnerable place. The EXP transition from the vulnerable place to the vulnerability-free place is allowed to be enable if there is at least one token in the vulnerable place. Similarly, the immediate transition from the vulnerable place to the vulnerability-free place is enable when the vulnerable place has at least one token and the dummy place has one token. Although we omit to show Vulnerability module of Uemura's model with multilevel vulnerable states, it can be built as the MRSPN where the EXP transition from the vulnerable place to the vulnerability-free place is removed.

All the GEN transitions are given by deterministic distributions, i.e., all the GEN transitions fire with constant time intervals. Model parameters are given in Tables 10 through 12. In the tables, the type column indicates transition types,

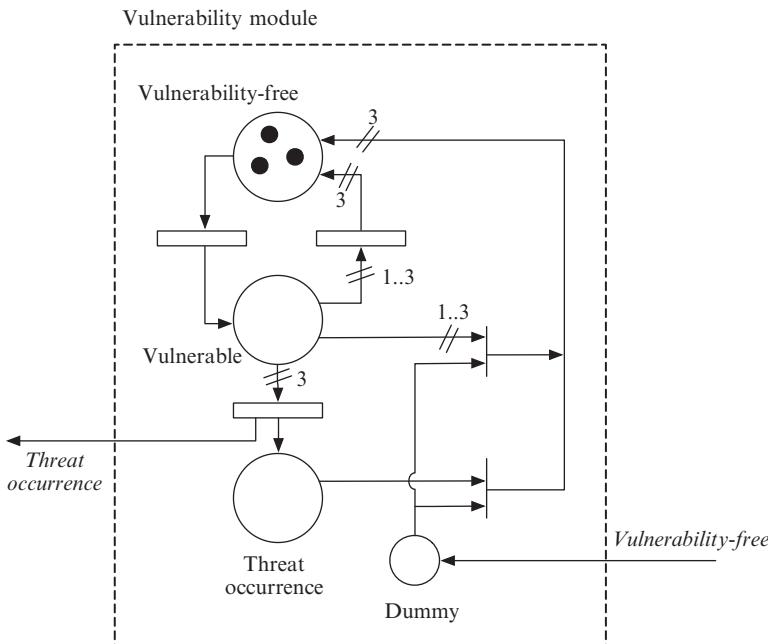


Fig. 17 Vulnerability module (type 1 with multilevel vulnerable states).

Table 10 Vulnerability module parameters

Type	From	To	Value
EXP	Vulnerability-free	Vulnerable	4.0
EXP	Vulnerable (only for type 1)	Vulnerability-free	0.72
EXP	Vulnerable	Threat occurrence	4.0

Table 11 Intrusion tolerance module parameters

Type	From	To	Value
EXP	Intrusion detection	Detection failure	0.13
EXP	Intrusion detection	Detection completion	0.083
IMM	Detection completion	Masking	0.5
IMM	Detection completion	Triage	0.5
EXP	Triage	Failure	0.17
GEN	Triage	Evaluation completion	0.08
IMM	Evaluation completion	Fail-secure	0.5
IMM	Evaluation completion	Graceful degradation	0.5
GEN	Detection failure	–	0.24
GEN	Masking	–	0.12
GEN	Failure	–	0.48
GEN	Fail-secure	–	40
GEN	Graceful degradation	–	30

Table 12 Maintenance module parameters

Type	From	To	Value
GEN	Set up	Maintenance	0.10
GEN	Maintenance	—	0.10

i.e., immediate transition (IMM), EXP transition (EXP), and GEN transition (GEN). The columns “from” and “to” present the input and output places which the corresponding transition is connected to. The last column “value” gives a transition rate for EXP, a weight for IMM or a deterministic transition time for GEN.

Figs. 18 and 19 illustrate transient behavior of point availabilities in Madan et al.’s and Uemura et al.’s models with multilevel vulnerable states. There figures well present the difference of transient behavior of point availabilities for Madan et al.’s and Uemura et al.’s models. From the results, we find the instantaneous time points that the system availability degrades by implementing a patch management. Such observations could not be found by the steady-state availability in [22]. This is a good insight by using MRSPN-based models with a transient analysis tool.

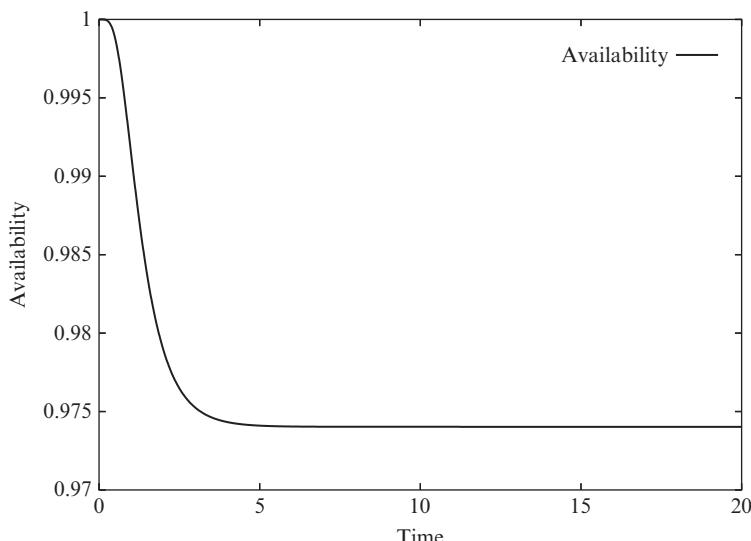


Fig. 18 Point availability for Madan et al.’s SITAR-security model with multilevel vulnerable states.

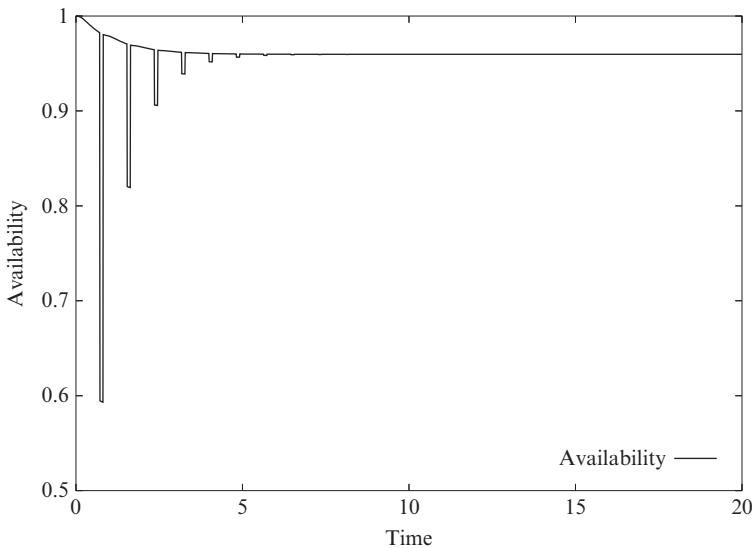


Fig. 19 Point availability for Uemura et al.’s SITAR-security model with multilevel vulnerable states.

6 Conclusions

In this chapter, we have presented availability models of an intrusion tolerant system with a control parameter called the preventive patch management time in SITAR. According to the semi-Markov modeling, not only the optimal preventive patch management policies but also the MTTSF have been analytically derived. We have also investigated quantitative effects of preventive maintenance based on the security patch releases in numerical examples. The lesson learned from the numerical examples was that the preventive patch release could improve the system availability and MTTSF effectively. The combination between the intrusion tolerance architecture and the preventive maintenance through monitoring the state was effective in some cases. On the other hand, for the MRSPN modeling, we have presented the MRSPNs for SITAR. MRSPNs give some usable methodologies for steady-state and transient analyses through some reliable tools [28, 29]. In fact, the numerical experiment has exhibited the transient analysis for the security measure with and without patch management under the multilevel vulnerable states.

The intrusion tolerant system is one of the attractive system architecture to prevent the security issues. In recent studies, the architecture of intrusion tolerant system with virtual machines has been proposed [30]. This is a promising way as a low cost architecture of intrusion tolerant system. In future, we will present the quantitative security evaluation for such a new architecture of intrusion tolerant system.

References

- [1] A. Avizienis, J.C. Laprie, B. Randell, C. Landwehr, Basic concepts and taxonomy of dependable and secure computing, *IEEE Trans. Dependable Secure Comput.* 1 (1) (2004) 11–33.
- [2] Y. Deswarte, L. Blain, J.C. Fabre, Intrusion tolerance in distributed computing systems, in: Proceedings of 1991 IEEE Symposium on Research in Security and Privacy, IEEE Computer Society Press, Silver Spring, MD, 1991, pp. 110–121.
- [3] P. Liu, J. Jing, P. Luenam, Y. Wang, L. Li, S. Ingsriswang, The design and implementation of a self-healing database system, *J. Intell. Inform. Syst.* 23 (3) (2004) 247–269.
- [4] P.E. Verissimo, N.F. Neves, M. Correia, Intrusion-tolerant architectures: concepts and design, in: R. Lemos, C. Gacek, A. Romanovsky (Eds.), *Architecting Dependable Systems*, LNCS, vol. 2677, Springer-Verlag, Berlin, 2003, pp. 3–36.
- [5] P.E. Verissimo, N.F. Neves, C. Cachin, J. Poritz, D. Powell, Y. Deswarte, R. Stroud, I. Welch, Intrusion-tolerant middleware, *IEEE Secur. Priv.* 4 (4) (2006) 54–62.
- [6] V. Gupta, V. Lam, H.V. Ramasamy, W.H. Sanders, S. Singh, Dependability and performance evaluation of intrusion-tolerant server architectures, in: Latin-American Symposium on Dependable Computing 2003, LNCS, vol. 2847, Springer-Verlag, Berlin, 2003, pp. 81–101.
- [7] R. Stroud, I. Welch, J. Warne, P. Ryan, A qualitative analysis of the intrusion-tolerant capabilities of the MAFTIA architecture, in: Proceedings of 34th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2004), IEEE Computer Society Press, Silver Spring, MD, 2004, pp. 453–461.
- [8] Y. Djemaiel, S. Rekhis, N. Boudriga, Intrusion detection and tolerance: a global scheme, *Int. J. Commun. Syst.* 21 (2) (2007) 211–230.
- [9] M. Haridasan, R. Renesse, SecureStream: an intrusion-tolerant protocol for live-streaming dissemination, *Comput. Commun.* 31 (2008) 563–575.
- [10] Y. Deswarte, D. Powell, Internet security: an intrusion-tolerance approach, *Proc. IEEE* 94 (2) (2006) 432–441.
- [11] B. Littlewood, S. Brocklehurst, N. Fenton, P. Mellor, S. Page, D. Wright, J. Doboson, J. McDermid, D. Gollmann, Towards operational measures of computer security, *J. Comput. Secur.* 2 (2) (1993) 211–229.
- [12] E. Jonsson, T. Olovsson, A quantitative model of the security intrusion process based on attacker behavior, *IEEE Trans. Softw. Eng.* 23 (4) (1997) 235–245.
- [13] R. Ortalo, Y. Deswarte, M. Kaaniche, Experimenting with quantitative evaluation tools for monitoring operational security, *IEEE Trans. Softw. Eng.* 25 (5) (1999) 633–650.
- [14] T. Uemura, T. Dohi, Quantitative evaluation of intrusion tolerant systems subject to DoS attacks via semi-Markov cost models, in: M.K. Denko, C.-S. Shih, K.-C. Li, S.-L. Tsao, Q.-A. Zeng, S.-H. Park, Y.-B. Ko, S.-H. Hung, J.-H. Park (Eds.), *Emerging Directions in Embedded and Ubiquitous Computing*, LNCS, 4809 Springer-Verlag, Berlin, 2007, pp. 31–42.
- [15] T. Uemura, T. Dohi, Optimizing security measures in an intrusion tolerant database system, in: T. Nanya, F. Maruyama, A. Pataricza, M. Malek (Eds.), *Service Availability: 5th International Service Availability Symposium*, LNCS, 5017 Springer-Verlag, Berlin, 2008, pp. 26–42.
- [16] H. Wang, P. Liu, Modeling and evaluating the survivability of an intrusion tolerant database system, in: D. Gollmann, J. Meier, A. Sabelfeld (Eds.), *European Symposium on*

- Research in Computer Security 2006, LNCS, vol. 4189, Springer-Verlag, Berlin, 2006, pp. 207–224.
- [17] J. Zheng, H. Okamura, T. Dohi, Survivability analysis of VM-based intrusion tolerant systems, *IEICE Trans. Inf. Syst.* E98-D (12) (2015) 2082–2090.
 - [18] J. Zheng, H. Okamura, T. Dohi, Mean time to security failure of VM-based intrusion tolerant systems, in: Proceedings of the 15th International Workshop on Assurance in Distributed Systems and Networks (ADSN2016), 2016.
 - [19] F. Wang, F. Gong, C. Sargor, K. Goseva-Popstojanova, K.S. Trivedi, F. Jou, SITAR: a scalable intrusion-tolerant architecture for distributed services, in: Proceedings of 2nd Annual IEEE Systems, Man and Cybernetics, Information Assurance Workshop, West Point, NY, 2001.
 - [20] B.B. Madan, K. Goseva-Popstojanova, K. Vaidyanathan, K.S. Trivedi, Modeling and quantification of security attributes of software systems, in: Proceedings of the 32nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2002), IEEE Computer Society Press, Silver Spring, MD, 2002, pp. 505–514.
 - [21] B.B. Madan, K. Goseva-Popstojanova, K. Vaidyanathan, K.S. Trivedi, A method for modeling and quantifying the security attributes of intrusion tolerant systems, *Perform. Eval.* 56 (1) (2004) 167–186.
 - [22] T. Uemura, T. Dohi, N. Kaio, Availability modeling of an intrusion tolerant system with preventive maintenance, in: S.-H. Sheu, T. Dohi (Eds.), *Advanced Reliability Modeling III—Global Aspect of Reliability and Maintainability*, McGraw-Hill, New York, 2008, pp. 655–662.
 - [23] R. Fujimoto, H. Okamura, T. Dohi, Security evaluation of an intrusion tolerant system with MRSPNs, in: Proceedings of 4th International Conference on Availability, Reliability and Security (ARES 2009), IEEE Computer Society Press, Silver Spring, MD, 2009, pp. 427–432.
 - [24] H. Choi, V.G. Kulkarni, K.S. Trivedi, Markov regenerative stochastic Petri nets, *Perform. Eval.* 20 (1994) 337–357.
 - [25] A. Bobbio, S. Garg, M. Gribaudo, A. Horváth, M. Sereno, M. Telek, Modelling software systems with rejuvenation restoration and checkpointing through fluid stochastic Petri nets, in: Proceedings of 6th International Workshop on Petri Nets and Performance Models, IEEE Computer Society Press, Silver Spring, MD, 1999, pp. 82–91.
 - [26] R. German, C. Lindemann, Analysis of stochastic Petri nets by the method of supplementary variables, *Perform. Eval.* 20 (1994) 317–335.
 - [27] M. Telek, A. Horváth, Transient analysis of Age-MRSPNs by the method of supplementary variables, *Perform. Eval.* 45 (2001) 205–221.
 - [28] R. German, C. Kelling, A. Zimmermann, G. Hommel, TimeNET: a tool for evaluating non-Markovian stochastic Petri nets, *Perform. Eval.* 24 (1995) 69–87.
 - [29] C. Lindemann, DSPNexpress: a software package for the efficient solution of deterministic and stochastic Petri nets, *Perform. Eval.* 22 (1995) 3–21.
 - [30] J. Lau, L. Barreto, J.D.S. Fraga, An infrastructure based in virtualization for intrusion tolerant services, in: 2012 IEEE 19th International Conference on Web Services (ICWS), 2012, pp. 170–177.

Mathematics and acoustics

3

P. Gulia, A. Gupta

Indian Institute of Technology Mandi, Kamand, India

The word “acoustic” is come from a Greek word “ακουστικός,” which means “ready to hear.” Acoustic engineering has a lot of application in noise control and audio industries. Acoustic is the study of sound and its propagation is either in pipe, cavity, free space, or closed space. Sound is omnipresent in our life. Each of us is a source of sound. When we talk or do some activity, it produces sound. In air, chirping of birds produce sound; in ocean, waves produce sound. Sound propagates as mechanical waves. Sound is a form of mechanical vibration. When we talk, we can feel some vibration in our throat. Sound needs a medium to propagate in which medium particles vibrate back and forth parallel to the propagation of wave. Due to vibration of particles, wave particles somewhere compress together called “compression of wave” and somewhere spread apart called “rarefaction of wave.” Compression of wave creates high pressure region and rarefaction creates low pressure region. Sound waves have consecutive high pressure and low pressure regions. So, sound propagates as a pressure disturbance. Acoustics represents sound wave in mathematical form using some fundamental laws. There are many acoustic applications; some are used to enhance the sound and some are used to reduce noise. In this chapter, horn, pipes, waveguides, and periodic structure are explained elaborately. First we need to explain a few terms which are used in this chapter.

- *Sound* is a form of vibration that propagates in terms of mechanical wave. Sound needs a medium to propagate. This is the difference between sound and electromagnetic wave. Electromagnetic wave does not require a medium to propagate.
- *Frequency* is the number of cycles per unit time and represented by f .
- *Wavelength* is the distance traveled by wave in one complete cycle and represented by λ .

$$\lambda = \frac{c}{f}$$

where c is the speed of sound.

- *Wave vector* is the measure of wave and represented by k . It shows the direction of propagation of wave and its magnitude is represented by:

$$k = \frac{2\pi}{\lambda} = \frac{2\pi f}{c} = \frac{\omega}{c}$$

where ω is the natural frequency.

- *Sound pressure level (SPL)* is the level which tells how much pressure is applied in terms of sound with respect to a reference pressure. And reference pressure is 20 μPa .

- *Decibel scale (dB)* converts the sound pressure level into logarithmic scale.

$$dB = 10 \log_{10} \left(\frac{p}{p_{ref}} \right)$$

$$p_{ref} = 20 \text{ } \mu\text{Pa}$$

- *Sound power level (SWL)* is the level which tells how much power is radiated in terms of sound with respect to a reference power. And reference power is 10^{-12} W .

$$B = 10 \log_{10} \left(\frac{W}{W_{ref}} \right)$$

$$W \text{ is sound power; } W = \frac{p^2}{2\rho c}.$$

Let P be an instantaneous pressure of sound wave.

P_0 is equilibrium pressure $= 1.013 \times 10^5 \text{ Pa}$

p is acoustic pressure and $p = P - P_0$

ρ_0 is equilibrium density

ρ is instantaneous density and $\rho = \rho_0 + s\rho_0$

s is called condensation and $s = \frac{\rho - \rho_0}{\rho_0}$

β is bulk modulus and $p = \beta s$

c is speed of sound and $c^2 = \frac{\beta}{\rho}$

1 Linear Euler equation

Let us take a small fluid element of volume $dx dy dz$ having mass dm and acceleration \vec{a} as shown in Fig. 1.

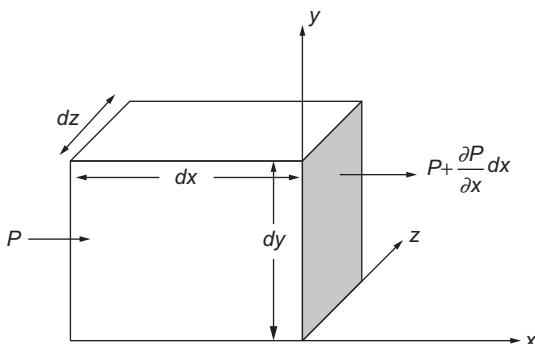


Fig. 1 An elemental fixed volume having fluid flow showing pressure into and out of the volume in x -direction.

Net force in x -direction is

$$df_x = \left[P - \left(P + \frac{\partial P}{\partial x} dx \right) \right] dy dz = -\frac{\partial P}{\partial x} dx dy dz$$

Similarly net force in y - and z -directions can be calculated.

Net force on fluid element includes force in x -, y -, z -directions with gravitational field in vertical direction.

$$df = -\left(\frac{\partial P}{\partial x} + \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} \right) dx dy dz + g\rho dV = -\nabla P dV + g\rho dV$$

$$dm \vec{a} = -\nabla P dV + g\rho dV \quad (1)$$

$$\vec{a} = \frac{\partial \vec{u}}{\partial t} + u_x \frac{\partial \vec{u}}{\partial x} + u_y \frac{\partial \vec{u}}{\partial y} + u_z \frac{\partial \vec{u}}{\partial z} = \frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} \quad (2)$$

Combining Eqs. (1), (2)

$$\begin{aligned} -\nabla P + g\rho &= \rho \left(\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} \right) \\ P &= P_0 + p \end{aligned} \quad (3)$$

$$\nabla P = \nabla P_0 + \nabla p = \rho_0 g + \nabla p \text{ and } \rho = \rho_0 + s\rho_0$$

Putting the value of ∇P and ρ in Eq. (3)

$$-\frac{1}{\rho_0} \nabla p + gs = (1+s) \left(\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} \right) \quad (4)$$

“ s ” is a condensation and a very small quantity. So neglecting s and assuming that $\frac{\partial \vec{u}}{\partial t} \gg (\vec{u} \cdot \nabla) \vec{u}$, Eq. (4) will be

$$\rho_0 \frac{\partial \vec{u}}{\partial t} = -\nabla p \quad (5)$$

This is called Euler equation of wave.

2 Wave equation

Continuity equation of wave [1] is

$$\rho_0 \frac{\partial s}{\partial t} + \nabla \cdot (\rho_0 \vec{u}) = 0$$

Differentiating with respect to t

$$\rho_0 \frac{\partial^2 s}{\partial t^2} + \nabla \cdot \left(\rho_0 \frac{\partial \vec{u}}{\partial t} \right) = 0 \quad (6)$$

Taking divergence of Eq. (5)

$$\nabla \cdot \left(\rho_0 \frac{\partial \vec{u}}{\partial t} \right) = -\nabla \cdot \nabla p = -\nabla^2 p \quad (7)$$

Combining Eqs. (6), (7)

$$\rho_0 \frac{\partial^2 s}{\partial t^2} - \nabla^2 p = 0 \quad (8)$$

and $p = \beta s$;

$$\rho_0 \frac{\partial^2 s}{\partial t^2} = \frac{\rho_0}{\beta} \frac{\partial^2 p}{\partial t^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}$$

So Eq. (8) will be

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (9)$$

This equation is called linear lossless wave equation for sound propagation in fluid with phase speed c .

When a wave is traveling in a medium, some of the sound energy reflects back after colliding with medium particles. It means that when a wave is traveling forward in a medium, some of its part gets reflected back and travels in backward direction. So, total sound pressure at that boundary will be a combination of forward traveling and backward traveling waves. Pressure function of wave traveling in x -direction is represented by

$$p = p^+ + p^- = A e^{i(\omega t - kx)} + B e^{i(\omega t + kx)}$$

A and B are constants. p^+ stands for forward traveling wave and p^- stands for backward traveling wave.

Applying the Euler equation

$$\rho_0 \frac{\partial \vec{u}}{\partial t} = ikA e^{i(\omega t - kx)} - ikB e^{i(\omega t + kx)}$$

Integrating with respect to t

$$\rho_0 \vec{u} = \frac{k}{\omega} (A e^{i(\omega t - kx)} - B e^{i(\omega t + kx)})$$

$$\vec{u} = \frac{p^+}{\rho_0 c} - \frac{p^-}{\rho_0 c} = u^+ - u^-$$

\vec{u} is the particle velocity of wave. So, resultant velocity of wave particle is the combination of velocity of forward traveling and backward traveling waves.

3 Sound propagation in horn

Horn is like a tapered waveguide for sound and this shape of waveguide supports impedance matching between sound source and free air which increase the amplitude of sound. Horns are used to increase the efficiency of sound propagation from source to free space. Cornet and Trumpet are some applications of horn. Gramophone was also designed by using the concept of horn more than hundred years ago. Horns have a small opening as an inlet and gradually diverging area as an outlet. So horns act on a principle just opposite to human ear. In human ear, outer ear which receives sound is broader than the area of inner ear [2]. There are many applications of horn-like loudspeaker horn, horn antenna, vehicle and musical horns, etc.

- *Loudspeaker horn* is a device which uses horn to increase the efficiency of loudspeaker. In throat of horn, waves have high acoustic pressure with very small displacement due to smaller area of throat. But at open end of horn, it has larger area and waves have low pressure with high displacement. So pressure varies from large to small pressure in horn. Ratio of diaphragm area and throat area is called compression ratio. High compression ratio makes the horn more efficient to couple with diaphragm and enhance its efficiency to vibrate diaphragm with large amplitude [3].
- *Musical horn* [4] is an instrument in which one end has horn-like shape with a large area and other end at throat is extended like tube and wounded like coil. This coil has small holes, which help to control the pitch of the sound propagating inside instrument.
- *Vehicle horn* is a device used in automobiles to show the vehicle's presence or to warn other vehicles.
- *Ear trumpet* is designed for those people who face difficulties in hearing. Ear trumpet is designed like a small horn, which collects sound from surroundings and increases sound pressure at throat. So ear trumpet increases the amplitude of sound inside ear.
- *Horn antenna* [5] is sometimes also called microwave horn used to collect and transmit the radio wave in space. It is a horn shaped waveguide to collect radio waves and to make them propagating like beam. Horn antenna can be used for large bandwidth (Fig. 2).

To derive the equation for horn, we assume that area of horn varies very slowly meaning that it diverges very slowly and there is no sudden increase in area. Horns are assumed of infinite length to avoid the effect of reflected wave.

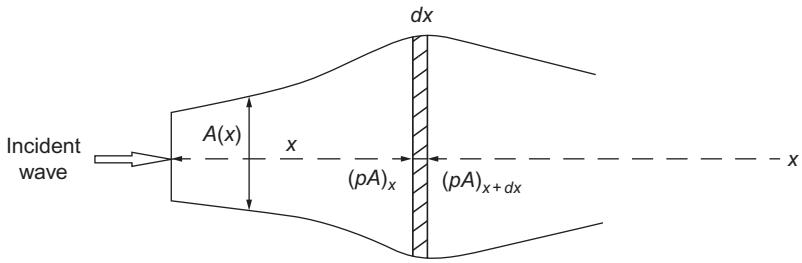


Fig. 2 A horn showing pressure force on a small elemental strip.

Let a small strip of area dA and length dx at a distance x from inlet. Left boundary acts as a planar wave front. Wave is propagating in x -direction, v is particle velocity of wave, ρ_0 is density of medium, c is speed of sound in medium, and k is wave vector.

Let $p = p_0 e^{i(\omega t - kx)}$

$$\frac{\partial^2 p}{\partial t^2} = -\omega^2 p_0 e^{i(\omega t - kx)} = -\omega^2 p \quad (10)$$

Force balance

$$\begin{aligned} (pA)_x - (pA)_{x+\delta x} + p dA &= \rho_0 A dx \frac{\partial v}{\partial t} \\ pA - (pA + \partial p A) + p dA &= \rho_0 A dx \frac{\partial v}{\partial t} \\ -\frac{\partial(pA)}{\partial x} + p \frac{\partial A}{\partial x} &= \rho_0 A \frac{\partial v}{\partial t} \\ -\left(A \frac{\partial p}{\partial x} + p \frac{\partial A}{\partial x}\right) + p \frac{\partial A}{\partial x} &= \rho_0 A \frac{\partial v}{\partial t} \\ -A \frac{\partial p}{\partial x} &= \rho_0 \frac{\partial(Av)}{\partial t} \\ -\frac{\partial}{\partial x} \left(A \frac{\partial p}{\partial x}\right) &= \rho_0 \frac{\partial^2(Av)}{\partial x \partial t} \end{aligned} \quad (11)$$

Mass balance

$$\begin{aligned} (\rho_0 Av)_x - (\rho_0 Av)_{x+\delta x} &= A dx \frac{\partial \rho}{\partial t} \\ \rho_0 Av - (\rho_0 Av + \rho_0 \partial Av) &= A dx \frac{\partial \rho}{\partial t} \\ \rho_0 \left(\frac{\partial Av}{\partial x} \right) &= -\frac{A}{c^2} \frac{\partial p}{\partial t} \quad \text{where } \frac{\partial p}{\partial \rho} = c^2 \\ \rho_0 \left(\frac{\partial^2 Av}{\partial x \partial t} \right) &= -\frac{A}{c^2} \frac{\partial^2 p}{\partial t^2} \end{aligned} \quad (12)$$

Combining Eqs. (10)–(12)

$$\begin{aligned}\frac{\partial}{\partial x} \left(A \frac{\partial p}{\partial x} \right) &= \frac{A}{c^2} (-\omega^2 p) \\ \frac{\partial^2 p}{\partial x^2} + \frac{1}{A} \frac{\partial A}{\partial x} \frac{\partial p}{\partial x} + \frac{\omega^2}{c^2} p &= 0 \\ \frac{d^2 p}{dx^2} + \frac{d(\ln A)}{dx} \frac{dp}{dx} + \frac{\omega^2}{c^2} p &= 0\end{aligned}\tag{13}$$

Here A is the function of x . This equation was given by A.G. Webster and equation is named as Webster's horn equation.

For constant area duct, $\ln A$ will be 0.

$$\frac{d^2 p}{dx^2} + \frac{\omega^2}{c^2} p = 0\tag{14}$$

This equation is called the Helmholtz equation of wave.

Horns are categorized mainly in two types

- Conical horn
- Exponential horn

3.1 Conical horn

Conical horns are widely used as horn antenna in radar guns and microwave radiometers to direct radio waves in a beam (Fig. 3).

Area function of conical horn is represented by $A(x) = Ax^2$. Using Webster horn equation,

$$\frac{d^2 p}{dx^2} + \frac{d(\ln A(x))}{dx} \frac{dp}{dx} + \frac{\omega^2}{c^2} p = 0$$

Let $p = p_0 e^{i(\omega t - kx)}$

$$\frac{\partial p}{\partial x} = -ip_0 k e^{i(\omega t - kx)}$$

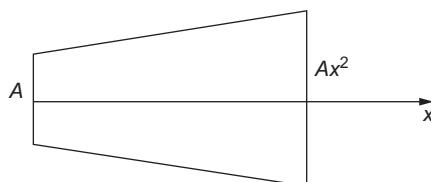


Fig. 3 Conical horn.

$$\frac{\partial^2 p}{\partial x^2} = -p_0 k^2 e^{i(\omega t - kx)}$$

$$\frac{\partial^2 p}{\partial t^2} = -\omega^2 p_0 e^{i(\omega t - kx)}$$

$$\frac{\partial A}{\partial x} = 2Ax$$

Putting above values into Webster horn equation:

$$-p_0 k^2 e^{i(\omega t - kx)} + \frac{1}{Ax^2} 2Ax(-i)p_0 k e^{i(\omega t - kx)} = -\frac{1}{c^2} \omega^2 p_0 e^{i(\omega t - kx)}$$

$$k^2 + \frac{2ik}{x} - \frac{\omega^2}{c^2} = 0$$

$$k = -\frac{i}{x} + \sqrt{\omega^2 - \frac{c^2}{x^2}} = k'_c - \frac{i}{x}$$

$$\text{where } k'_c = \sqrt{\omega^2 - \frac{c^2}{x^2}}.$$

$$p = p_0 e^{i(\omega t - kx)} = p_0 e^{i(\omega t - k'_c x + i)} = p_0 e^{-1} e^{(\omega t - k'_c x)}$$

$$\text{For positive value of } k'_c: \omega^2 > \frac{c^2}{x^2}$$

$$\omega_c = \frac{c}{x}$$

ω_c is called the cut-off frequency of conical horn. Below this frequency, sound cannot propagate through a conical horn.

3.2 Exponential horn

Area function of exponential horn is defined by $A(x) = A_0 e^{2\epsilon x}$, where ϵ is flare coefficient. Sometimes it is also called spherical wave horn [6] (Fig. 4).

$$\text{Here } \frac{\partial A}{\partial x} = 2\epsilon A_0 e^{2\epsilon x}.$$

Now using Webster horn equation

$$-p_0 k^2 e^{i(\omega t - kx)} + \frac{1}{A_0 e^{2\epsilon x}} 2\epsilon A_0 e^{2\epsilon x} (-i)p_0 k e^{i(\omega t - kx)} = -\frac{1}{c^2} \omega^2 p_0 e^{i(\omega t - kx)}$$

$$k^2 + 2\epsilon ik - \frac{\omega^2}{c^2} = 0$$

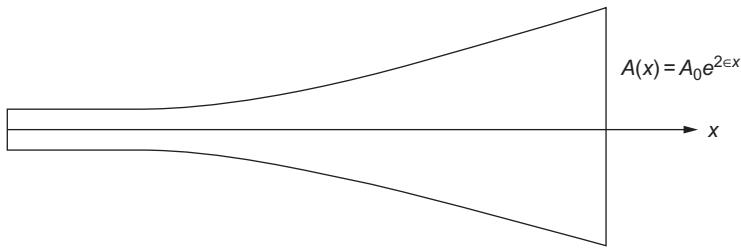


Fig. 4 Exponential horn.

$$k = -\epsilon i + \frac{\sqrt{\omega^2 - (\epsilon c)^2}}{c} = k'_e - \epsilon i$$

$$\text{where } k'_e = \frac{\sqrt{\omega^2 - (\epsilon c)^2}}{c}.$$

$$p = p_0 e^{i(\omega t - kx)} = p_0 e^{i(\omega t - k'_e x + \epsilon ix)} = p_0 e^{-\epsilon x} e^{(\omega t - k'_e x)}$$

For real value of k'_e : $\omega > c \epsilon$

$$\omega_c = c \epsilon$$

$c \epsilon$ is cut-off frequency of exponential horn. Below this frequency, sound cannot propagate through an exponential horn. Therefore, exponential horn can also be used like high pass filter because it cannot allow the frequency lower than the cut-off frequency.

$$p = p_0 e^{-\epsilon x} e^{(\omega t - k'_e x)}$$

Term $e^{-\epsilon x}$ shows that sound wave is decaying exponentially with distance; this is called evanescent nature of sound wave.

Speed of sound in exponential horn

$$c' = \frac{\omega}{k'_e} = \frac{\omega}{\sqrt{\left(\frac{\omega}{c}\right)^2 - \epsilon^2}} = \frac{c}{\sqrt{1 - \left(\frac{\omega}{\omega_c}\right)^2}}$$

$$c' = \frac{c}{\sqrt{1 - \left(\frac{\omega}{\omega_c}\right)^2}} = f(\omega)$$

Exponential horn is said to have a dispersive medium because speed of sound in exponential horn is a function of frequency.

4 Sound propagation in pipes

The way of sound propagation in pipe depends on the shape and size of the pipe; it also depends on whether the pipe is open ended or close ended. Resonance in pipe occurs when the frequency of sound wave matches with its natural frequency. At resonance, sound waves show a standing wave pattern. Distance between two consecutive nodes or antinodes is measured by half wavelength of sound.

This concept is used in Kundt's tube [7] to calculate speed of sound in fluid. Kundt's tube is a horizontal tube made with transparent glass. One end of this tube is fixed and the other end is having a piston that can be adjusted at desired length. This tube contains a very fine powdery material like talc or dust. When a signal at particular frequency is generated at one end of the tube, the other end of the tube is blocked by a piston. Piston is adjusted so that there is a sudden increase in the sound in the tube. Sudden increase in the sound indicates resonance in the tube. Length of the path traveled by the sound wave from piston to fixed end and fixed end to piston is measured by wavelength. At resonance, dust particles are equally spaced at nodes, and distance between these nodes is half of the wavelength of sound. By multiplying the frequency with wavelength, the speed of the sound can be calculated.

$$c = \lambda f$$

where λ is the wavelength and f is the frequency of the sound wave.

Let L be the pipe length (Fig. 5). If wavelength of the sound wave is very longer than the diameter of the pipe, then the wave behaves as a plane wave.

Incident pressure p will be the combination of forward traveling wave and backward traveling wave.

$$\begin{aligned} p &= Ae^{i(\omega t - kx)} + Be^{i(\omega t + kx)} \\ p &= e^{i\omega t}[(A + B)\cos kx + i(B - A)\sin kx] \end{aligned} \tag{15}$$

Particle velocity is a vector quantity and can be represented by:

$$\begin{aligned} v &= \frac{p^+}{\rho_0 c} - \frac{p^-}{\rho_0 c} = \frac{A}{\rho_0 c}e^{i(\omega t - kx)} - \frac{B}{\rho_0 c}e^{i(\omega t + kx)} \\ v &= \frac{e^{i\omega t}}{\rho_0 c}[(A - B)\cos kx - i(A + B)\sin kx] \end{aligned} \tag{16}$$



$x=0$

$x=L$

Fig. 5 A pipe of length L showing sound propagation from $x=0$ to $x=L$.

Acoustic impedance Z is defined as

$$Z = \frac{p}{v} = \frac{\rho_0 c [(A+B) \cos kx + i(B-A) \sin kx]}{[(A-B) \cos kx - i(A+B) \sin kx]}$$

$$\frac{Z}{r} = \frac{[(A+B) \cos kx + i(B-A) \sin kx]}{[(A-B) \cos kx - i(A+B) \sin kx]}$$

where $r = \rho_0 c$ = characteristic impedance.

At $x=0$

$$p[0] = A + B$$

$$rv[0] = A - B$$

$$p = p[0] \cos kx - irv[0] \sin kx$$

$$rv = rv[0] \cos kx - ip[0] \sin kx$$

These equations are called transmission line equations.

At $x=L$

$$\begin{aligned} \frac{Z[L]}{r} &= \frac{p[0] \cos kL - irv[0] \sin kL}{rv[0] \cos kL - ip[0] \sin kL} \\ \frac{Z[L]}{r} &= \frac{\frac{p[0]}{v[0]} - ir \frac{\sin kL}{\cos kL}}{r - i \frac{p[0] \sin kL}{v[0] \cos kL}} \end{aligned} \quad (17)$$

$$\frac{Z[L]}{r} = \frac{Z[0] - ir \tan kL}{r - iZ[0] \tan kL}$$

$$z[0] = \frac{rZ[L] + ir^2 \tan kL}{r + iZ[L] \tan kL} \quad (18)$$

Case I: If $kL = \frac{\pi}{2}$

$$k = \text{wave vector} = \frac{2\pi}{\lambda} = \frac{\omega}{c}$$

where λ is wavelength of wave.

$$kL = \frac{2\pi}{\lambda} = \frac{\pi}{2}$$

$$L = \frac{\lambda}{4}$$

$$\frac{Z[L]}{r} = \frac{\frac{Z[0]}{\tan kL} - ir}{\frac{r}{\tan kL} - iZ[0]} = \frac{r^2}{Z[0]}$$

Case II: If we have rigid inlet boundary.

Rigid boundary does not allow the transmitting sound so acoustic impedance of rigid boundary is infinite.

$$Z[0] \rightarrow 0$$

$$\frac{Z[L]}{r} = \frac{1 - ir \frac{\tan kL}{Z[0]}}{\frac{r}{Z[0]} - i \tan kL} = i \cot kL$$

If $kL \ll 1$ then $\cos kL \sim 1$ and $\sin kL \sim kL$

$$Z[L] = \frac{ir}{kL}$$

4.1 Piston with fixed end

Let a piston with one driving end and one fixed end. At fixed end, pressure is maximum and velocity is zero. No sound can transmit beyond this boundary. So, acoustic impedance is infinite at fixed boundary (Fig. 6).

When

$$Z[L] = \infty$$

$$Z[0] = \frac{r + ir^2 \frac{\tan kL}{Z[L]}}{\frac{r}{Z[L]} + i \tan kL} = \frac{r}{i \tan kL}$$



Fig. 6 A pipe having piston at $x=0$ with one end fixed at $x=L$.

For resonance, reactance part of the impedance is zero.

$$Z[0] = \frac{-ir}{\tan kL} = 0$$

$$\tan kL = \infty; \quad kL = (2n - 1)\frac{\pi}{2} \quad \text{where } n = 1, 2, 3, \dots$$

$$\frac{2\pi}{\lambda}L = (2n - 1)\frac{\pi}{2}$$

$$L = (2n - 1)\frac{\lambda}{4}; \quad \frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}, \frac{7\lambda}{4}, \dots$$

$$\frac{2\pi f}{c}L = (2n - 1)\frac{\pi}{2}$$

$$f = (2n - 1)\frac{c}{4L}; \quad \frac{c}{4L}, \frac{3c}{4L}, \frac{5c}{4L}, \frac{7c}{4L}, \dots$$

Therefore, resonance frequencies of piston with fixed end are odd multiples of fundamental frequency. These are called odd harmonics.

At resonance frequency, waves show a standing wave pattern inside pipe [8]. When sound waves propagate in pipe, interference of wave occurs due to reflection of sound waves. Constructive interference results in maximum sound pressure and a point called antinode. Destructive interference results in zero pressure and a point called node. Fig. 7 shows a standing wave pattern of sound waves in pipe for first, third, and fifth harmonics. Fundamental frequency or first harmonic frequency is the lowest frequency on which sound waves show a standing wave pattern.

4.2 Piston with open end

Now the piston end is open. At open end, pressure is zero. So, acoustic impedance is zero at open end (Fig. 8).

$$Z[L] = 0$$

$$Z[0] = ir \tan kL$$

For resonance:

$$\tan kL = 0; \quad kL = n\pi$$

$$\frac{2\pi}{\lambda}L = n\pi$$

$$L = \frac{n\lambda}{2}; \quad \frac{\lambda}{2}, \frac{2\lambda}{2}, \frac{3\lambda}{2}, \dots$$

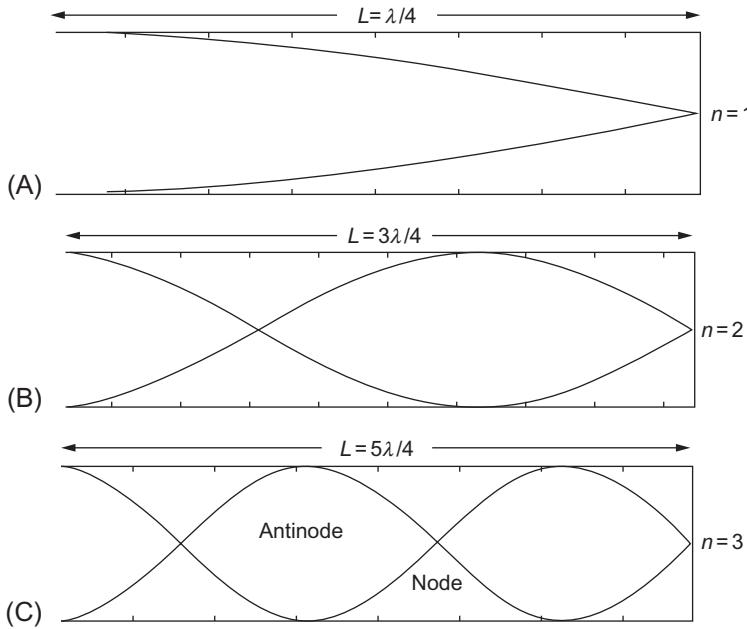


Fig. 7 Displacement standing wave pattern for odd harmonic with frequency (A) $c/4L$, (B) $3c/4L$, and (C) $5c/4L$.

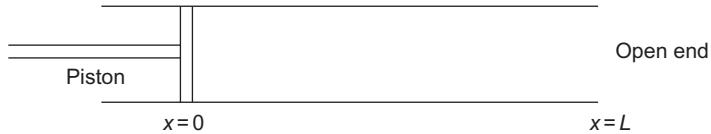


Fig. 8 A pipe having piston at $x=0$ and open end at $x=L$.

$$\frac{2\pi f}{c}L = n\pi$$

$$f = \frac{nc}{2L}; \quad \frac{c}{2L}, \frac{2c}{2L}, \frac{3c}{2L} \dots$$

Resonance frequency is even multiple of fundamental frequency in case of a piston with open end. These are two cases of sound propagation in pipe with fixed and open end. When open end is flanged then there will be some end correction in resonance frequency (Fig. 9).

$$f_{flanged} = \frac{nc}{2(L + 0.85a)}$$

where $n = 1, 2, 3, \dots$ and a is pipe diameter.

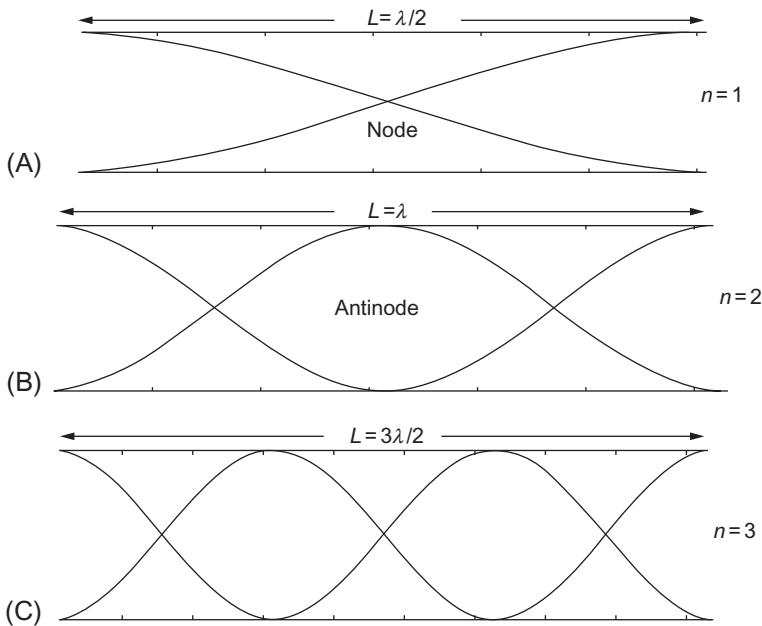


Fig. 9 Displacement standing wave pattern for open-ended pipe at frequency (A) $c/2L$, (B) c/L , and (C) $3c/2L$.

5 Diffraction, scattering, reflection, and refraction of sound

- *Diffraction* occurs for all types of waves [9]. When wave travels across barriers, the ability of wave to change its direction or to bend is called diffraction. If there is a large post between you and sound source, then you can hear only when the wavelength of sound wave is large enough so that they can bend around post [10]. If you are going to a music show and you are standing outside at the small opening of the door, you can hear music due to small opening diffraction of sound wave. But if you are at a large distance between the source and the post, then you cannot sense the presence of post by nature of sound. Diffraction helps sound wave to bend around the post or any obstacle. With the increase in the wavelength of the sound wave, the amount of diffraction increases and vice versa. Many birds like owls can communicate for a longer distance due to longer wavelength of their hoots which diffracts around trees and covers a longer distance. In a recent research, it is found that elephants emit low frequency infrasonic waves for communication to each other for long distances.
- *Scattering* [9] refers to collision between the sound wave and the rough surface. This is the only difference between diffraction and scattering. When sound waves collide with an obstacle or barrier, it diffracts but if waves collide with a rough surface or a medium having randomly distributed obstacles, each obstacle deflects some amount of the energy out of an incident wave. When wave is scattered only by one scattering center, it is called single scattering. But when wave is scattered by multiple number of scattering center, it is called

multiple scattering. If a surface shows a very strong scattering, then intensity I of scattered sound is given by Lambert's cosine law [9].

$$I(r, \theta) = B(1 - \alpha) \frac{\cos \theta}{r^2} dS$$

Here dS is the elemental area of rough surface with absorption coefficient α , B is the energy incident per unit area per unit second, r is the distance between the sound source and the rough surface, and θ is the angle subtended by chosen direction of scattering of wave to the normal of surface.

- *Reflection* [10] refers to change in the direction of the wave front at the interface of two media and after changing direction, sound waves come back in the medium from which it is originated. When sound waves propagate in a medium to another medium, if it is not absorbed or not transmitted, then it will be fully reflected. Law of reflection for sound waves is same as that for light wave which states that angle of incidence of wave is equal to angle of reflection. Incident sound waves are forward traveling waves where reflected sound waves are backward traveling waves. In a closed pipe, interference phenomenon occurs due to reflection of wave which results formation of nodes (destructive interference) and anti-nodes (constructive interference). So, near the rigid surface, sound intensity is almost double to its original amplitude due to construction interference of wave. Reflection of sound is the main reason for producing echo and reverberation inside the room.
- *Refraction* [10] refers to bending of the sound wave when entering from one medium to the other. When sound wave gets refracted, its speed is also changed. Refraction follows the Snell's law which states that ratio of sine of angle of incidence and angle of refraction is always constant (Fig. 10).

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{c_1}{c_2} = \frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1}$$

Here θ_1 and θ_2 are the angle of incidence and angle of refraction, respectively; c is the speed of sound; λ is the wavelength of sound wave; and n is the refractive index of medium. Subscripts 1 and 2 show medium 1 and medium 2.

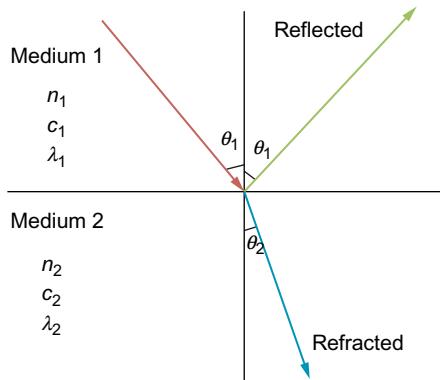


Fig. 10 Sound wave propagates from one medium to another, reflected wave bends in same medium with same angle, and refracted wave bends with different angle in 2nd medium.

6 Sound propagation in cavity and waveguide

As the name suggests, waveguide guides the sound waves. Waveguide restricts the free space propagation of sound wave in one direction. Waveguide looks like a large hollow cavity with one or more sides open. Application of the waveguide is used in sound propagation in atmosphere or ocean, digital delay lines, folded horn loudspeakers, etc.

6.1 Sound propagation in cavity

Let us take a living room with dimensions l_x, l_y, l_z . The room will act like a cubical cavity. It is assumed that walls of cavity are rigid and particle velocity across the wall of cavity is zero (Fig. 11).

Using the Euler's equation

$$-\nabla p = 0$$

So

$$\left(\frac{\partial p}{\partial x}\right)_{x=0, l_x} = \left(\frac{\partial p}{\partial y}\right)_{y=0, l_y} = \left(\frac{\partial p}{\partial z}\right)_{z=0, l_z} = 0 \quad (19)$$

When walls of a room are rigid, sound wave collides with walls and gets reflected and again collides and gets reflected. Due to the repeated reflections, interference will occur which will result in standing wave pattern in this close cavity.

Inside the cavity, wave is free to move in all three directions. So, pressure function for waves propagating inside the cavity will be

$$p(x, y, z, t) = p_x p_y p_z e^{i\omega t} \quad (20)$$

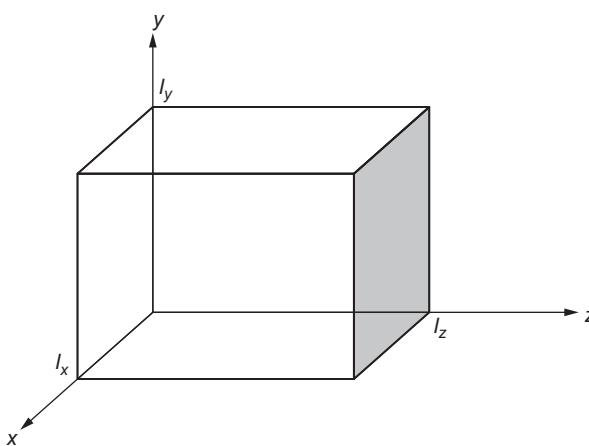


Fig. 11 A small living room having dimensions l_x, l_y, l_z in x -, y -, and z -directions, respectively.

Using wave equation in three dimensional space

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}$$

$$\frac{\partial^2 p_x}{\partial x^2} p_y p_z e^{i\omega t} + \frac{\partial^2 p_y}{\partial y^2} p_z p_x e^{i\omega t} + \frac{\partial^2 p_z}{\partial z^2} p_x p_y e^{i\omega t} = \frac{1}{c^2} (-\omega^2) p_x p_y p_z e^{i\omega t}$$

$$\frac{\partial^2 p_x}{\partial x^2} \frac{1}{p_x} + \frac{\partial^2 p_y}{\partial y^2} \frac{1}{p_y} + \frac{\partial^2 p_z}{\partial z^2} \frac{1}{p_z} + \frac{\omega^2}{c^2} = 0$$

By using separation of variables method

$$\left(\frac{\partial^2 p_x}{\partial x^2} + k_x^2 p_x = 0 \right)$$

$$\left(\frac{\partial^2 p_y}{\partial y^2} + k_y^2 p_y = 0 \right)$$

$$\left(\frac{\partial^2 p_z}{\partial z^2} + k_z^2 p_z = 0 \right)$$

Using boundary condition Eq. (19), solution of Eq. (20) becomes

$$p_{lmn} = A_{lmn} \cos(k_{xl}x) \cos(k_{ym}y) \cos(k_{zn}z) = \frac{1}{8} A_{lmn} \sum_{\pm} e^{i(\omega_{lmn}t \pm k_{xl}x \pm k_{ym}y \pm k_{zn}z)} \quad (21)$$

At rigid walls, particle velocity is zero and impedance is infinite.

$$\text{And } \tan(k_{xl}l_x) = 0; k_{xl} = \frac{l\pi}{l_x} \text{ where } l = 0, 1, 2, \dots$$

$$k_{ym} = \frac{m\pi}{l_y} \text{ where } m = 0, 1, 2, \dots$$

$$k_{zn} = \frac{n\pi}{l_z} \text{ where } n = 0, 1, 2, \dots$$

l, m, n are the modes of wave propagation.

Resultant angular frequency of cavity is

$$\omega^2 = c^2 \left(k_{xl}^2 + k_{ym}^2 + k_{zn}^2 \right)$$

Allowed frequency in cavity is [11]

$$(\omega_{lmn})_{\text{allowed}} = c \sqrt{\left(\frac{l_x \pi}{L_x}\right)^2 + \left(\frac{m \pi}{L_y}\right)^2 + \left(\frac{n \pi}{L_z}\right)^2}$$

At different mode (l, m, n) , frequency will be different. The modes, at which allowed frequencies have same value, are called degenerate modes.

6.2 Sound propagation in waveguide

A waveguide as the name suggests, guides the wave in a particular direction. Ear canal and stethoscopes are examples of waveguide. Waveguide is like a duct which contains some medium. It behaves like a transmission line for wave propagation. There are some applications of waveguide eg. air conditioner duct, car muffler, and ultrasound imaging.

Let us take a waveguide having uniform rectangular cross-section in $x-y$ plane (Fig. 12). Sound is propagating in z -direction. It is assumed that the side walls of waveguide are rigid and sound source is present at left boundary ($z=0$) of waveguide. When sound propagates through waveguide, wave will show the standing wave pattern due to reflection from rigid walls.

So, pressure in the waveguide can be represented using Eq. (21)

$$p_{lm} = A_{lm} \cos(k_{xl}x) \cos(k_{ym}y) e^{i(\omega t - k_z z)} \quad (22)$$

Using wave equation

$$\begin{aligned} \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} &= \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \\ -p_{lm}(k_{xl}^2 + k_{ym}^2 + k_z^2) &= -p_{lm} \left(\frac{\omega}{c}\right)^2 \\ k_z &= \sqrt{\left(\frac{\omega}{c}\right)^2 - (k_{xl}^2 + k_{ym}^2)} = \sqrt{\left(\frac{\omega}{c}\right)^2 - k_{lm}^2} \end{aligned} \quad (23)$$

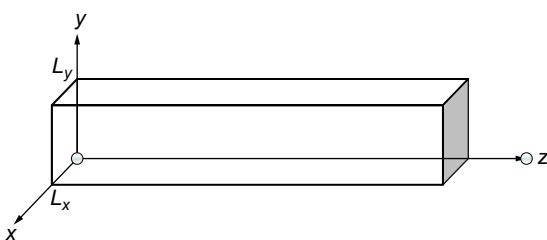


Fig. 12 A waveguide having constant rigid cross-section $L_x L_y$, and sound wave propagates along z -direction.

For k_z to be real, $\omega \geq c k_{lm}$, and if $\omega < c k_{lm}$, k_z will be a complex number. Let k_r and k_i be real and imaginary parts of k_z .

$$k_z = \pm i \left(\sqrt{k_{lm}^2 - \left(\frac{\omega}{c}\right)^2} \right)$$

+ sign can be eliminated because wave always shows an evanescent behavior. Amplitude of sound always decays and finally dies out after a certain distance.

So, the Eq. (22) can be rewritten as

$$\begin{aligned} p_{lm} &= A_{lm} \cos(k_{xl}x) \cos(k_{ym}y) e^{i(\omega t + (\sqrt{k_{lm}^2 - (\frac{\omega}{c})^2})z)} \\ p_{lm} &= A_{lm} e^{-(\sqrt{k_{lm}^2 - (\frac{\omega}{c})^2})z} \cos(k_{xl}x) \cos(k_{ym}y) e^{i(\omega t - k_r z)} \end{aligned} \quad (24)$$

Exponential term represents that standing waves have an evanescent behavior that decay exponentially with distance.

When $\omega = c k_{lm}$, frequency is called cut-off frequency of waveguide. k_z will be zero at this frequency and below this frequency, there is no sound propagation in waveguide.

$$\omega_{cutoff} = ck_{lm}$$

Phase speed of a mode in waveguide is given by

$$c_p = \frac{\omega}{k_z} = \frac{\omega}{\sqrt{\left(\frac{\omega}{c}\right)^2 - k_{lm}^2}} = \frac{c}{\sqrt{1 - \left(\frac{\omega_{cutoff}}{\omega}\right)^2}} \quad (25)$$

Let k be a wave vector at angle θ to the k_z (Fig. 13).

$$\cos \theta = \frac{k_z}{k} = \sqrt{1 - \left(\frac{\omega_{cutoff}}{\omega}\right)^2}$$

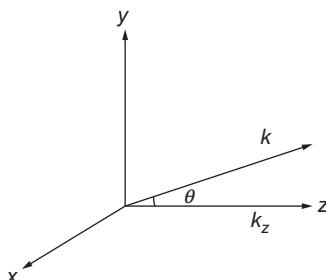


Fig. 13 Wave vector in spatial coordinate has angle θ with k_z .

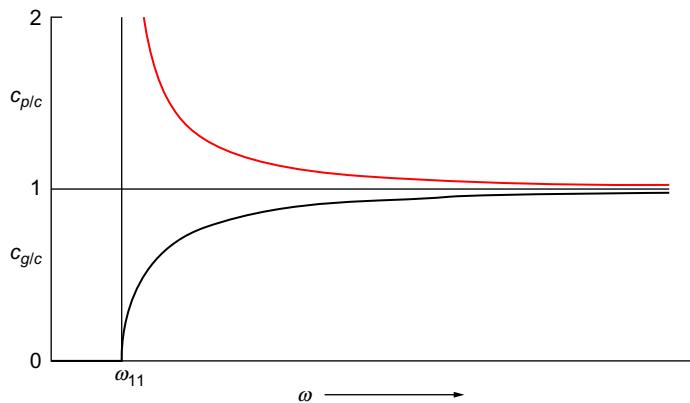


Fig. 14 Phase speed and group speed for mode (1, 1) versus natural frequency.

From Eq. (25)

$$c_p = \frac{c}{\cos \theta}$$

Group speed is the speed at which energy is transmitted perpendicular to the waveguide or in the direction of propagation of wave.

$$c_g = c \cos \theta$$

A graph between phase speed and group speed versus frequency at mode (1, 1) is represented in Fig. 14 which shows that group speed is started at zero and approaches to “ c ” at higher frequency whereas phase speed has higher value at lower frequency and approaches to “ c ” at lower frequency.

7 Sound propagation in periodic structure (sonic crystal)

Periodic structures are artificial structures in which scatterers are arranged periodically in a medium. Phononic and sonic crystals are the examples for acoustic periodic structures. Periodic structures are used as sound attenuator, sound diffuser, frequency filter, liquid sensors, etc. Destructive interference is the cause of sound attenuation in particular range of frequency called band gap. In band gap, sound does not propagate. A full band gap is the range of frequency in which propagation of sound is not possible in all directions of spatial coordinate. First acoustic experiment on periodic structure was done by Martinz Sala in 1995. The structure had a long hollow cylinder as scatterers and these scatterers were arranged with a regular period of distance on a large circular domain. Sound source was present at one side of the structure and the receiver was present at the other side. He observed that other side of periodic structure sound pressure level is 20 dB less than the source side.

Sonic and phononic crystals are both periodic structures. In sonic crystals, scatterers are made of solid material and arranged in a fluidic material like air, water, etc. In phononic crystals, scatterers are made of solid or fluid and arranged in a solid medium. When scatterers are arranged along one direction, these are called 1-D periodic structures. When scatterers are arranged periodically in two and three directions, these are called 2-D and 3-D periodic structures. Fig. 15 shows a 2-D sonic crystal in which circular scatterers are arranged in two directions.

Scatterers act as a sound hard in sonic crystal and sound wave gets reflected back after colliding with the scatterers. Nowadays, resonant cavity is used inspired by Helmholtz resonator to get full band gap. Resonant cavity looks like a spherical container having a small hole. Below the resonance frequency, sound cannot propagate in the structures made of resonators. So resonators are also used for frequency filters. Main purpose of using periodic structures is to reduce the sound transmission. Sound attenuation is measured in terms of insertion loss. Insertion loss is the difference between sound pressure level without inserting periodic structure and sound pressure level with periodic structure.

$$IL = SPL_{\text{without structure}} - SPL_{\text{with structure}} = 20 \log_{10} \frac{p_i}{p_o}$$

Here p_i is the inlet wave pressure and p_o is the pressure of the outgoing wave.

Wave propagation in a 2-D periodic structure is represented by a Helmholtz equation.

$$\frac{d^2 p}{dx^2} + \frac{\omega^2}{c^2} p = 0 \quad \text{where } p = f(x, y)$$

Sound waves are harmonic waves that change their amplitudes with time and distance. Sound pressure is the combination of forward and backward traveling waves and is a function of time and distance

$$p(x, t) = p_i e^{i(\omega t - kx)} + p_r e^{i(\omega t + kx)} = e^{i\omega t} (p_i e^{-ikx} + p_r e^{ikx}) = e^{i\omega t} p(x) \quad (26)$$

where $p(x) = p_i e^{-ikx} + p_r e^{ikx}$

p_i and p_r stand for incident and reflected wave pressures.

$$u(x, t) = \frac{1}{\rho c} e^{i\omega t} (p_i e^{-ikx} - p_r e^{ikx}) = e^{i\omega t} U(x) \quad (27)$$

where $U(x) = \frac{1}{\rho c} (p_i e^{-ikx} - p_r e^{ikx})$

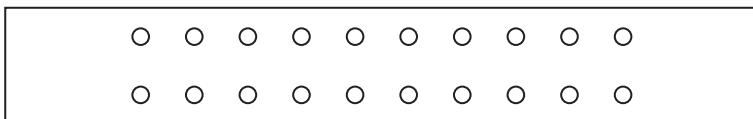


Fig. 15 A sonic crystal having circular scatterers arranged periodically in 2-D on a rectangular domain.

At $x=0$

$$p_{x=0} = p_i + p_r \quad (28)$$

$$u_{x=0} = \frac{p_i}{\rho c} - \frac{p_r}{\rho c} \quad (29)$$

By solving above two equations p_i and p_r can be calculated.

By similar approach, $p_{x=l}$ can be calculated and one can find the transmission loss through periodic structure by subtracting the pressure at inlet and outlet [12].

If sound source is radiating a wave having pressure p_i , the sound power at inlet is determined by using the following equation:

$$W_{in} = \frac{p_{in}^2}{2\rho c}$$

where c is the speed of the sound and ρ is the density of the host material.

If absolute sound pressure at outlet is p_{abs} , then sound power level at outlet is

$$W_{in} = \frac{p_{abs}^2}{2\rho c}$$

$$\text{Net transmission Loss} = 10 \times \log_{10} \left(\frac{W_{in}}{W_{out}} \right).$$

Above calculation is done for a finite periodic structure. To study the band gap in the case of infinite periodic structures is also possible by using Bloch–Floquet theorem. Bloch–Floquet theorem is applied only for a periodic structure and a unit cell is sufficient to analyze the infinite periodic structure. When a 2-D periodic structure has periodicity “ a ” along length and periodicity “ b ” along breath, then pressure function at the boundary of unit cell is described by following equation [13].

$$p(\vec{r} + \vec{a} + \vec{b}) = p(x) \exp[i(k_x a + k_y b)] \quad (30)$$

where \vec{r} is a position vector and k_x and k_y are wave vectors in x - and y -directions, respectively.

When wave vector is swept over a Brillouin zone, k_x ranges from (0 to π/a) and k_y ranges from (0 to π/b), a full band gap occurs in a particular range of frequency. Fig. 16 shows six eigenmodes on different frequencies when k_x sweeps from 0 to π/a . Lattice constant a is 10 cm and structure is periodic only in one direction. Gray region shows band gap and no eigenmode can propagate within this frequency range.

8 Muffler

Muffler is also called a silencer. It is used to reduce the noise which is emitted by exhaust of internal combustion engine. Mufflers are used in automobiles to reduce the amplitude of sound pressure which is created by engine. There are mainly

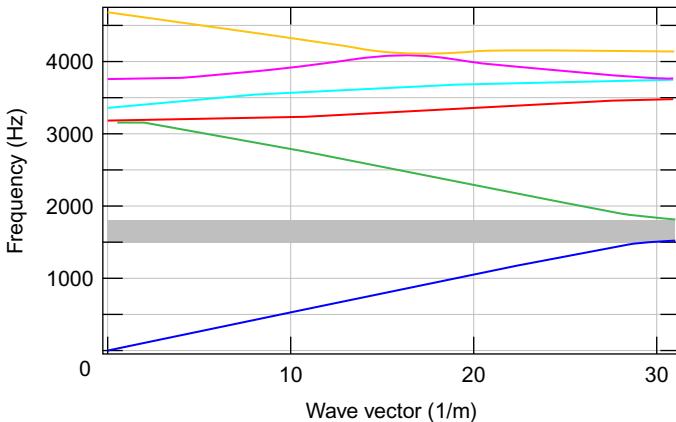


Fig. 16 A dispersion analysis of six eigenmodes when kx sweeps from 0 to p_i/a and ky is 0.

two types of mufflers present in the market. One is the absorber muffler and the other one is the reflector muffler. In the absorber type muffler, silencer is covered by some absorbing material. When engine noise passes through the muffler, absorbing material absorbs some amount of noise and reduces its amplitude. This design is quite simple and helps to maintain low back pressure.

In reflector type muffler, destructive interference is the cause of sound reduction in the muffler. There is some set of tube in which some holes are drilled inside the muffler. These tubes are designed in such a way that wave can reflect again and again after colliding with tubes so that destructive interference can partially cancel some of the waves. Holes are drilled so that sound wave can pass through the tubes.

A diesel engine radiates more exhaust noise in comparison to a gasoline engine [14]. A diesel engine without muffler can radiate exhaust noise ranges from 100 to 125 dB where a gasoline engine without muffler can radiate exhaust noise ranges from 90 to 100 dB. An engine with muffler can produce noise 15–35 dB less than the engine without muffler. Muffler is a very important part of engine that is used to reduce the exhaust sound transmission frequently.

8.1 Transmission loss in mufflers

Let us take a simple geometry of muffler having inlet and outlet tube of area A_1 and the area of expansion chamber of muffler is A_2 shown in Fig. 17 [15].

P_i and P_o are the inlet and outlet pressures; P_c is the pressure in expansion chamber; Superscript + and – present the incident and reflected waves; ρ and c are the density and speed of sound of medium inside expansion chamber; k is the wave number; α is the expansion ratio; and $\alpha = \frac{A_2}{A_1}$.

At boundary $x=0$

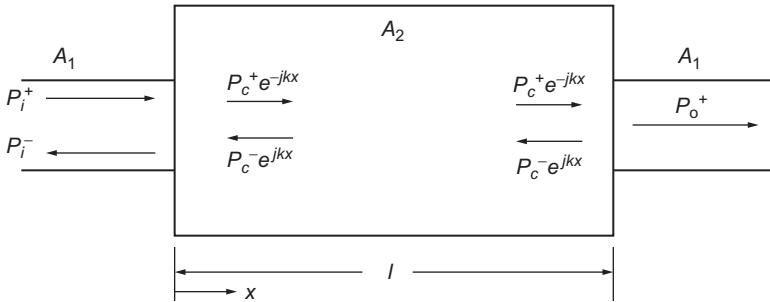


Fig. 17 Muffler with one-dimensional wave propagation.

Acoustic pressure:

$$\begin{aligned} P_i^+ e^{-jkx} + P_i^- e^{jkx} &= P_c^+ e^{-jkx} + P_c^- e^{jkx} \\ P_i^+ + P_i^- &= P_c^+ + P_c^- \end{aligned} \quad (31)$$

Volume velocity:

$$\begin{aligned} A_1 \left(\frac{P_i^+}{\rho c} - \frac{P_i^-}{\rho c} \right) &= A_2 \left(\frac{P_c^+}{\rho c} - \frac{P_c^-}{\rho c} \right) \\ (P_i^+ - P_i^-) &= \alpha (P_c^+ - P_c^-) \end{aligned} \quad (32)$$

By solving Eqs. (31), (32)

$$P_i^+ = \frac{1+\alpha}{2} P_c^+ + \frac{1-\alpha}{2} P_c^- \quad (33)$$

At boundary $x=1$

Acoustic pressure:

$$\begin{aligned} P_c^+ e^{-jkx} + P_c^- e^{jkx} &= P_0^+ \\ P_c^+ e^{-jkl} + P_c^- e^{jkl} &= P_0^+ \end{aligned} \quad (34)$$

Volume velocity:

$$\begin{aligned} A_2 \left(\frac{P_c^+}{\rho c} - \frac{P_c^-}{\rho c} \right) &= A_1 \left(\frac{P_0^+}{\rho c} \right) \\ (P_c^+ - P_c^-) &= \frac{1}{\alpha} (P_0^+) \end{aligned} \quad (35)$$

Solving Eqs. (34), (35)

$$P_c^+ = \frac{\alpha+1}{2\alpha} P_0^+ e^{jkl} \quad (36)$$

$$P_c^+ = \frac{\alpha-1}{2\alpha} P_0^+ e^{-jkl} \quad (37)$$

Using Eqs. (33), (36), (37)

$$\begin{aligned}\frac{P_i^+}{P_0^+} &= \frac{(\alpha+1)^2}{4\alpha} e^{jkl} - \frac{(\alpha-1)^2}{4\alpha} e^{-jkl} \\ \frac{P_i^+}{P_0^+} &= \cos kl + j \left(\frac{1+\alpha^2}{2\alpha} \right) \sin kl\end{aligned}\tag{38}$$

Transmission loss in muffler is calculated as: $TL = 20 \log_{10} \frac{P_i^+}{P_0^+}$

$$TL = 20 \log_{10} \left(\cos kl + j \left(\frac{1+\alpha^2}{2\alpha} \right) \sin kl \right)\tag{39}$$

This is a very simple method to calculate transmission loss in mufflers. There are many other methods like transfer matrix method, two source method, and decomposition method to measure the transmission loss in mufflers [16,17].

- *Decomposition method*

$$TL = 20 \log_{10} \frac{P_i}{P_t} + 10 \log_{10} \frac{A_i}{A_o}$$

where P_i is the incident sound power and P_t is the transmitted sound power. A_i and A_o are the area of inlet and outlet of muffler.

- *Transfer matrix method*

$$TL = 10 \log_{10} \left(1 + \frac{1}{4} \left(\alpha - \frac{1}{\alpha} \right)^2 \sin kl \right)$$

This chapter describes the basic equations to understand the propagation of sound wave in horns, pipes, waveguides, periodic structures, and mufflers. There are many applications of acoustics like acoustic lenses, noise barriers, sound diffusers, acoustic imaging, and ear muffs. Basic understanding of acoustic engineering helps to understand the mechanism of these applications.

9 Other acoustic applications

- *Acoustic lens* works in a same way like an optical lens. Optical lens is used to focus light rays and acoustic lens is used to focus sound waves. At the time of World War II, acoustic lenses were used to detect the aircraft of enemy by collecting the sound of their engine. Sound waves focus on a point after striking to acoustic lens and an intense sound wave is created which acts as a sound bullet and used for medical applications.
- *Noise barriers* are used to keep traffic noise low or to protect the human beings who are living in a very noisy area from loud noise. Noise barrier cannot completely erase the noise but can reduce the noise up to half of its amplitude. Noise barriers are like solid vertical walls of height up to eight meters. Noise barrier reduce the noise by absorbing some of the waves

and reflect the other waves back to take a longer path over barrier. Waves have evanescent behavior. Thus amplitude of sound waves reduces while traveling over a long distance.

- *Sound diffusers* are used to diffuse or spread the sound evenly in a room to improve its acoustic performance. Sound diffusers diminish the intensity of the sound wave by dispersing it over a large surface area. A room with high diffusion and low echo effect is considered as an ideal room. Diffusion refers to the distribution of energy equally in a given space. In a perfect diffusive room, value of reverberation time is equal while listener changes his position in the room. Reverberation time is the time taken by sound to decay 60 dB from its initial level. Reverberation time is represented by T_{60} .

$$T_{60} = 0.161 \frac{V}{S\alpha}$$

Here V is the volume of the room whose reverberation time has to be calculated; S is the total surface area of room; and α is the average sound absorption coefficient of the room surface.

Sound diffusers are like an acoustic panel which helps to spread the sound evenly in the given environment. Sound diffusers are mainly used in recording or music studio, control rooms, etc.

- *Acoustic imaging* is used in medical application like ultrasound or lithotripsy. A very high frequency wave is sent to the particular body part which reflected back after striking the bone or cells of the body part. Sensors are used to capture the clear image of the body part using reflected waves. Shipping industry uses sonar to detect or to collect information about shoals of fishes in sea and gynecologists use ultrasound to image the baby in mother's womb.
- *Ear muffs* are used to protect human ears from loud noise. A very loud noise can harm our ear drum. Ear muffs are made of some absorbing materials and used to protect ears of human beings who are working in a very noise environment.
- *Musical instruments* include guitar, trumpet, violin, etc. When sound wave of two different frequencies interacts, then mixing of these two frequencies creates high and low air pressures, which produce a new sound.

10 Conclusion

This chapter helps to understand the basics of sound wave propagating in different structures mathematically. Some devices are used to enhance the noise like horns and some devices are used to attenuate the sound like periodic structure. Acoustic structure plays a very important role in protecting the building from earthquakes and for sound insulation. There are also some medical applications like ultrasound, lithotripsy associated with the acoustic engineering. This chapter contains the mathematical derivations and formulation of sound propagation in pipes, waveguides, cavities, horns, and periodic structure and their applications include noise barriers, filters, mufflers, musical instruments, ear muffs, etc.

References

- [1] L.E. Kinsler, A.R. Frey, A.B. Coppens, J.V. Sanders, *Fundamentals of Acoustics*, fourth ed., John Wiley and Sons, New York, 2009.
- [2] J.A. Dearman, *The Book of Horns*, Avantgarde Acoustic, Germany, 2010.

- [3] D.A. Moore, United States Patent, US Patent No. US 7,506,721 B2, 2009.
- [4] J. Beakes, The horn parts in Handel's operas and oratorios and the horn players who performed in these works, Dissertation, City University of New York, 2007.
- [5] P.N. Kumar, M.D. Rao, D.G.S. Reddy, Design and testing of wide band flat scalar ring horn antenna system for ground station receiver applications, *Int. J. Adv. Res. Comput. Commun. Eng.* 4 (9) (2015) 140–144.
- [6] T.H.E. Purpose, O.F.A. Horn, Horn Theory: An Introduction, Part 1, Audio Express, 2008. Available from: <https://www.grc.com/acoustics/an-introduction-to-horn-theory.pdf>, 2008.
- [7] P.K. Jeethendra Kumar, Ajeya PadmaJeeth, K. Santhosh, Velocity of sound in air the Kundt's tube experiment, Lab exp., KamalJeeth Instrumentation and Service Unit 13 (1) (2013) 44–48.
- [8] S.W. Rienstra, A. Hirschberg, *An Introduction to Acoustics*, IWDE 92-06, Eindhoven University of Technology, Netherlands, 2016.
- [9] H. Kuttruff, *Acoustics: An introduction*, Taylor & Francis, London and New York, 2006.
- [10] Reflection, Refraction, Diffraction, and Wave Interference, World of Physics, Gale, 2001. Science in Context, <link.galegroup.com/apps/doc/CV2434500569/SCIC?u=albertak12&xid=683906c2>.
- [11] I. Ver, L. Beranek, *Noise and Vibration Control Engineering*, second ed., Wiley, Hoboken, New Jersey, 2005.
- [12] A. Gupta, K.M. Lim, C.H. Chew, A quasi two-dimensional model for sound attenuation by the sonic crystals, *J. Acoust. Soc. Am.* 132 (4) (2012) 2909–2914.
- [13] C. Kittel, *Introduction to Solid State Physics*, eighth ed., Wiley, California, 2005.
- [14] P.B.S. Patel, S. Khokhar, A review paper on design of muffler toward reduce sound in exhaust system of IC engine, in: Proceeding of International Conference on Emerging Trends in Engineering and Technology, ICETET-6, Nagpur, 2013, , pp. 565–570.
- [15] C.J. Wu, X.J. Wang, H.B. Tang, Transmission loss prediction on a single-inlet/double-outlet cylindrical expansion-chamber muffler by using the modal meshing approach, *Appl. Acoust.* 69 (2008) 173–178.
- [16] Z. Tao, A.F. Seybert, *A Review of Current Techniques for Measuring Muffler Transmission Loss*, Society of Automotive Engineering, Kentucky, 2003.
- [17] M.L. Munjal, *Acoustics of Ducts and Mufflers with Application to Exhaust and Ventilation System Design*, Wiley, New York, 1987.

Signature reliability of sliding window coherent system

4

A. Kumar, S.B. Singh

G. B. Pant University of Agriculture and Technology, Pantnagar, India

1 Introduction

Reliability theory is a sphere that connects mathematics and engineering. Reliability engineering is the regulation for protecting a system when the system works in a specified mode. Engineering systems have basically two states: working and failed. Failure of the system component forces one to study the impact of it on the system and subsequently on the reliability of the system. Yingkui and Jing [1] described the engineering system reliability in different ways. Out of several engineering systems, sliding window system (SWS) is one of applied system, widely in the multistate k -out-of- r -from n : G (F) system. Some other applications of SWS can be seen in the field such as flow control, quality control, service system, manufacturing, engineering system, communication and power transmission system, and computer and telephone networks. The SWS, which is introduced by Levitin [2], can have two possible states: working and failed. Habib et al. [3] studied a model consisting of n linearly ordered multistate elements and calculated the reliability of SWS. Levitin and Dai [4] have discussed a novel model, which generalizes the linear multistate SWS for multiple failures. Xiao et al. [5] determined the optimal load of elements in SWS and calculated the maximum reliability of the linear SWS.

Further, system signature is a very useful device to compare the system's characteristics. The signature is widely applied in the reliability economics, optimality criteria, and optimal reliability system designs. Samaniego [6] analyzed and estimated the coherent system signature reliability and expected the lifetime of the component having independent and identically distributed (i.i.d.) components. He discussed the system signature $b = (b_1, b_2, \dots, b_m)$ which depends on the probability function P_r , failure component l , and the number of components of the system m . If we have the above information then signature b_l of the l th smallest lifetime is given by

$$b_l = P_r(T = X_{l:m}), \quad l \in [m] \quad (1)$$

Further, the signature is the characteristic of the system design in which signature of the system depends only on structure function ϕ and lifetime components X . Boland [7] gave the method to calculate the signature of the system in terms of structure function as

$$b_l = \phi_{m-l+1} - \phi_{m-l} \quad (2)$$

where,

$$\phi_l = \frac{1}{\binom{m}{l}} \sum_{\substack{X \in \{0, 1\}^m \\ |X|=l}} \phi(X) \quad (3)$$

and $|X| = \sum_{i=1}^m x_i$.

The system signature reliability $\bar{G}_s(t)$ of i.i.d. element can be evaluated by Eq. (2) as

$$\bar{G}_s(t) = \sum_{l=1}^m (\phi_{m-l+1} - \phi_{m-l}) \bar{G}_{l:m}(t) \quad (4)$$

where, $\bar{G}_{l:m}(t) = P_r(X_{l:m} > t)$.

Owen [8,9] proposed the method to measure the degree of Barlow-Proschan index and signature of the coherent systems. Barlow and Proschan [10] introduced an index, which determined the degree for each element with the help of its structure function h . Boland [7] also discussed the degree of the i.i.d. components in the signature reliability on the basis of the structure function. The index is defined by m -tuple I^{BP} (Barlow-Proschan index), whose i th coordinate is the probability of the failure element l causing the system failure. The expression for $I^{BP(l)}$ is given by

$$I_{BP}^{(l)} = \sum_{h \subseteq [m]/\{l\}} \frac{1}{m \binom{m-1}{|h|}} (\phi(h \cup \{l\}) - \phi(h)) \quad (5)$$

Navarro et al. [11,12], Samaniego [13], Li and Zhang [14], Bhattacharya and Samaniego [15], and Bairamov and Arnold [16] also studied the system signature and minimum signature to compare their expected lifetime or mean time to failure (MTTF). Navarro and Rubio [17] determined the signature of the coherent system with i.i.d. elements. Navarro and Rychlik [18] also evaluated the expected lifetime of the system signature reliability. Marichal and Mathonet [19] calculated the signature reliability function and signature of the systems environment with the help of tail probability.

In this study the proposed algorithm evaluated the reliability of the SWS using extended universal moment generating function. The present study deals with structure, reliability function, and signature, which explains the features of the SWCS. In the present chapter we have considered an SWCS model which consists of n sliding windows and each sliding window comprises n parallel i.i.d. components. The study evaluates the different measures such as signature reliability, MTTF, Barlow-Proschan index using U -function technique, which has not been done earlier for the SWCS.

The remainder of the chapter is organized as follows: [Section 2](#) briefly introduces the evaluation of the SWCS. [Section 3](#) discusses the coherent system and expected lifetime. In [Section 4](#) a numerical example is taken for illustration and discussion of signature reliability, MTTF, Barlow-Proschman index, and expected cost rate. [Section 5](#) presents conclusions of the study.

2 Appraisal reliability of SWS

Let us consider an SWCS with k windows in which every component is connected in a parallel arrangement s_1, s_2, \dots, s_k , as shown in [Fig. 1](#). The considered SWCS is working if at least one component of each window is functioning. In the present chapter we propose to apply U -function given by Ushakov [20] to evaluate the reliability measures of the considered system.

The U -function can define all feasible states of the system components i associated with the probabilities of the state p_{qi} to the working state g_{qi} .

Then U -function of a system can be expressed as follows:

$$U(z) = \sum_{i=1}^2 p_{qi} z^{g_{qi}}, \quad q = 1, 2, \dots, k \quad (6)$$

where,

$$g_{qi} = \begin{cases} 1, & i = 1 \\ 0, & i = 2 \end{cases}$$

and

$$p_{qi} = \begin{cases} p_{qi}, & i = 1 \\ (1 - p_{qi}), & i = 2 \end{cases}$$

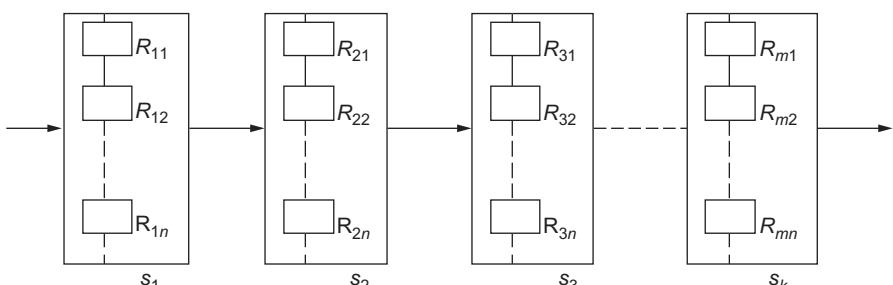


Fig. 1 Block diagram of SWCS model.

Applying the operator ϕ in the U -function of the components d and f (where d and f are the states of probabilities p_{qi}), we can have the U -function of SWCS as

$$\begin{aligned} U(z) &= \phi(U_d, U_f) \\ &= \sum_{m=1}^2 \sum_{q=1}^2 \left(\prod_{i=1}^2 p_{qiq} z^f(g_{qim}, g_{qim}) \right) \\ &= \sum_{m=1}^2 \sum_{q=1}^2 (p_{q1} \cdot p_{q2} z^{g_{nm}, g_{nm}}) \end{aligned} \quad (7)$$

Similarly, we have the general expression of the U -function for the anticipated model of k th components as

$$\begin{aligned} U(z) &= \phi(U_1(z), U_2(z), \dots, U_k(z)) \\ &= \sum_{j_1=1}^2 \sum_{j_1}^2 \cdots \sum_{j_k=1}^2 \left(\prod_{q=1}^k p_{qjq} z^f(g_{1j_1} g_{2j_2} \cdots g_{kj_k}) \right) \end{aligned} \quad (8)$$

2.1 Algorithm for estimation of reliability of the SWS

Step 1: Calculate the U -function of the particular component, allocate $F = 0$, $U_{1-r}(z) = z^{0, g_0}$.

Step 2: Replicate the value for $j = 1, 2, \dots, k$.

Step 3: Find $U_{j-r+1}(z) = U_{j-r}(z) \phi U_j(z)$.

Step 4: Eliminate all the terms which follow the condition $j \geq r$. Then add the values $\beta_f(U_{j-r+1}(z))$ to F .

Step 5: Obtain the reliability of SWS as $R = 1 - F$.

Proposition Let i and m be the i.i.d. components in a parallel arrangement where $i = 1, 2, \dots, k$, $m = 1, 2, \dots, n$, and P_1, P_2, \dots, P_K be the probability function of working components in the SWCS. Then reliability function R_i ($i = 1, 2, \dots, n$) of an SWCS consisting of n parallel components is given by

$$R_i = 1 - \prod_{m=1}^n (1 - R_{im}) \quad (9)$$

The expression of the structure function of the system with the help of Eq. (9) is given by

$$\phi(p_1, p_2, \dots, p_k) = p_1 \vee p_2 \vee \cdots \vee p_k$$

$$= \max_{1 \leq i \leq k} p_i$$

Hence,

$$R_1 = 1 - (1 - R_{11})(1 - R_{12}) \dots (1 - R_{1N}) \quad (10)$$

$$R_2 = 1 - (1 - R_{21})(1 - R_{22}) \dots (1 - R_{2N}) \quad (11)$$

⋮

$$R_K = 1 - (1 - R_{K1})(1 - R_{K2}) \dots (1 - R_{KN}) \quad (12)$$

From Eqs. (9)–(11), we get the reliability function of the parallel components as

$$R_i = 1 - \prod_{m=1}^n (1 - R_{im}), \quad i = 1, 2, \dots, k, \quad m = 1, 2, \dots, n \quad (13)$$

2.2 Algorithm for assessment of the signature of SWCS with its reliability function

Step 1: Compute the system signature of the structure function

$$B_l = \frac{1}{\binom{m}{m-l+1}} \sum_{\substack{h \subseteq [m] \\ |h|=m-l+1}} \phi(h) - \frac{1}{\binom{m}{m-l}} \sum_{\substack{h \subseteq [m] \\ |h|=m-l}} \phi(h) \quad (14)$$

Evaluate reliability polynomial of SWCS

$$h(p) = \sum_{i=1}^m c_i \binom{m}{i} p^i q^{n-i}$$

where,

$$c_j = \sum_{j=m-i+1}^m B_j, \quad i = 1, 2, \dots, m$$

Step 2: Calculate the tail signature of SWCS, i.e., $(m+1)$ -tuple $\bar{B} = (\bar{B}_0, \dots, \bar{B}_m)$ using

$$\bar{B}_l = \sum_{j=l+1}^m b_j = \frac{1}{\binom{m}{m-l}} \sum_{|H|=m-l} \phi(H) \quad (15)$$

Step 3: Compute the reliability function in the form of polynomial by using Taylor expansion about $y = 1$ by

$$p(y) = y^m h\left(\frac{1}{y}\right) \quad (16)$$

Step 4: Evaluate the tail signature of the SWCS reliability function with the help of Eq. (14) by

$$\bar{B}_l = \frac{(m-l)!}{m!} D^l p(1), \quad l = 0, \dots, m \quad (17)$$

Step 5: Calculate the signature of the SWCS using Eq. (16)

$$b = \bar{B}_{l-1} - \bar{B}_l, \quad l = 1, \dots, m \quad (18)$$

2.3 Algorithm to evaluate the expected lifetime of SWCS with minimum signature

Step 1: Calculate the expected lifetime of i.i.d. component SWCS which is exponentially distributed with mean ($\mu = 1$).

Step 2: Compute minimum signature SWCS with the expected lifetime of the reliability function by using

$$\bar{R}_T(t) = \sum_{j=1}^n c_j \bar{R}_{1:j}(t) = \sum_{j=1}^n d_j \bar{R}_{j:j}(t) \quad (19)$$

where, $\bar{R}_{1:j}(t) = P_r(Z_{1:j} > t)$ and $\bar{R}_{j:j}(t) = P_r(Z_{j:j} > t)$ for $j = 1, 2, \dots, n$.

Step 3: Evaluate $E(T)$ of SWCS of i.i.d. components by

$$E(T) = \mu \sum_{j=1}^n \frac{c_j}{j} \quad (20)$$

where, $c = (c_1, c_2, \dots, c_n)$ is a vector coefficient of minimal signature.

Step 4: Calculate the expected lifetime (MTTF) by reliability function $R(t)$ and probability density function $r(t)$ given by Høyland and Rausand [21] as

$$\frac{\partial}{\partial t} R(t) = -\frac{\partial}{\partial t} \ln R(t) = \frac{r(t)}{R(t)} = r(t) \quad (21)$$

2.4 Algorithm for estimation of the Barlow-Proschan index of SWCS

Calculate the Barlow-Proschan index of the i.i.d. components given by its structure function as

$$I_{BP}^{(i)} = \int_0^1 (\partial_i \hat{\phi})(y) dy = \int_0^1 (\partial_i h)(y) dy, \quad i = 1, 2, \dots, n \quad (22)$$

where $\hat{\phi}$ and h are structure and reliability functions of SWCS.

2.5 Algorithm for assessing the expected value of component X and expected cost rate of SWCS when working elements are failed

Step 1: Compute the number of failed elements at the time of system failure with signature

$$E(X) = \sum_{j=1}^n j \cdot b_j, \quad j = 1, 2, \dots, n \quad (23)$$

Step 2: Compute the $E(X)$ and $E(X)/E(T)$ of SWCS with minimum signature.

3 Coherent system and expected lifetime of the SWCS

A system is called coherent if its component is relevant and structure function ϕ is monotone. Consider a coherent system consisting of n components and let state x_i of i th element be defined as:

$$x_i = \begin{cases} 1, & \text{if } i\text{th element is working} \\ 0, & \text{if } i\text{th element is failed} \end{cases}$$

The structure function $\phi : \{0, 1\}^n \rightarrow \{0, 1\}$ is a plot that relates state vector X and the system is good if the value is 1 and down if the value is 0. Every coherent system satisfies the condition $\phi(0) = 0$ and $\phi(1) = 1$. The expected lifetime of the coherent system is evaluated when i.i.d. elements are exponentially distributed. The system components having average equal to one are computed using the minimal signature of the structure function. Finally, the reliability function $R(t)$ and the hazard rate function $r(t)$ of the system are given by

$$r(t) = h(t) = \frac{r(t)}{R(t)} \quad \text{and} \quad r(t) = \frac{r(t)}{\bar{R}(t)} = \lambda$$

$$\frac{\partial}{\partial t} R(t) = -\frac{\partial}{\partial t} \ln R(t) = \frac{r(t)}{\bar{R}(t)} = r(t)$$

$$\text{where, } R(t) = \int_0^t r(t) dt.$$

4 Illustration of the proposed model

Consider an SWCS in which every window has three parallel components with $n=4$, $r=2$, and $W=3$ as shown in Fig. 2. Each component can have either of the two states: working or failed. Let the performance rates of the element from 1 to 4 be 1, 2, 3, and 4, respectively.

Probability function P_i of the inner parallel system is defined as

$$P_i = 1 - \prod_{m=1}^n (1 - R_{im}) \tag{24}$$

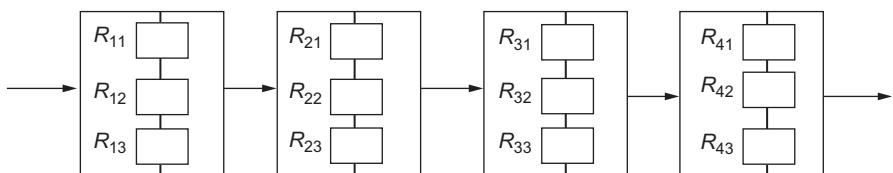


Fig. 2 SWCS with $n=4$, $r=2$, and $w=3$.

where $i = 1, 2, 3, 4$, $m = 1, 2, 3, 4$, and P_1, P_2, P_3, P_4 are the probability functions of the parallel systems given by

$$P_1 = 1 - \{(1 - R_{11})(1 - R_{12})(1 - R_{13})\} \quad (25)$$

$$P_2 = 1 - \{(1 - R_{21})(1 - R_{22})(1 - R_{23})\} \quad (26)$$

$$P_3 = 1 - \{(1 - R_{31})(1 - R_{32})(1 - R_{33})\} \quad (27)$$

$$P_4 = 1 - \{(1 - R_{41})(1 - R_{42})(1 - R_{43})\} \quad (28)$$

Now, the U -function of the SWCS is given by

$$U_i(z) = p_i z^i + (1 - p_i)z^0 \quad (29)$$

where $i = 1, 2, 3, 4$, P_i is the probability function, z^i is the performance rate, and z^0 is the nonperformance rate.

Hence, the U -functions of the SWCS components $U_i(z)$ ($i = 1, 2, 3, 4$) are given by

$$U_1(z) = p_1 z^1 + (1 - p_1)z^0 \quad (30)$$

$$U_2(z) = p_2 z^2 + (1 - p_2)z^0 \quad (31)$$

$$U_3(z) = p_3 z^3 + (1 - p_3)z^0 \quad (32)$$

$$U_4(z) = p_4 z^4 + (1 - p_4)z^0 \quad (33)$$

Using the Algorithm 2.1 of SWS, we get initial component of the U -function SWCS as
For $j = 1$

$$U_0(z) = \varphi(U_{-1}(z), U_1(z))$$

$$U_0(z) = \varphi\left(z^{0,(0,0)}, p_1 z^1 + (1 - p_1)z^0\right) \quad (34)$$

$$= p_1 z^{0,(0,1)} + (1 - p_1)z^{0,(0,0)}$$

For $j = 2$

$$U_1(z) = \varphi(U_0(z), U_2(z))$$

$$\begin{aligned} &= \varphi\left(p_1 z^{0,(0,1)} + (1 - p_1)z^{0,(0,0)}, p_2 z^2 + (1 - p_2)z^0\right) \\ &= p_1 p_2 z^{0,(1,2)} + p_1(1 - p_2)z^{0,(1,0)} + p_2(1 - p_1)z^{0,(0,2)} + (1 - p_1)(1 - p_2)z^{0,(0,0)} \end{aligned} \quad (35)$$

If performance rate r is greater than or equal to the given weight w , then by using condition $j \geq w$, unreliability F and $U_1(z)$ are expressed as

$$\begin{aligned} F &= p_1(1-p_2) + (1-p_1)(1-p_2) + p_2(1-p_1) \\ U_1(z) &= p_1 p_2 z^{0,(1,2)} \end{aligned} \quad (36)$$

For $j=3$

$$\begin{aligned} U_2(z) &= \varphi(U_1(z), U_3(z)) \\ &= \varphi(p_1 p_2 z^{0,(1,2)}, p_3 z^3 + (1-p_3)z^0) \\ &= p_1 p_2 p_3 z^{1,(2,3)} + p_1 p_2 (1-p_3) z^{1,(2,0)} \end{aligned} \quad (37)$$

Again, combining the alike terms F and $U_2(z)$ from Eqs. (36) and (37) respectively are expressed as

$$\begin{aligned} F &= p_1(1-p_2) + (1-p_1)(1-p_2) + p_2(1-p_1) + p_1 p_2 (1-p_3) \\ U_2(z) &= p_1 p_2 p_3 z^{1,(2,3)} \end{aligned} \quad (38)$$

For $j=4$

$$\begin{aligned} U_3(z) &= \varphi(U_2(z), U_4(z)) \\ &= \varphi(p_1 p_2 p_3 z^{1,(2,3)}, p_4 z^4 + (1-p_4)z^0) \\ &= p_1 p_2 p_3 p_4 z^{2,(3,4)} + p_1 p_2 p_3 (1-p_4) z^{2,(3,0)} \end{aligned} \quad (39)$$

$$F = p_1(1-p_2) + (1-p_1)(1-p_2) + p_2(1-p_1) + p_1 p_2 (1-p_3) \quad (40)$$

Hence, from Eqs. (38) and (40), we get the reliability R of the system as

$$R = p_1 p_2 p_3 \quad (41)$$

Hence, substituting the values of p_1 , p_2 , and p_3 in Eq. (41) from Eqs. (25), (26), and (27), we obtain the reliability function R of SWCS as

$$\begin{aligned} R &= (R_{11} + R_{12} + R_{13} - R_{11}R_{12} - R_{11}R_{13} - R_{12}R_{13} + R_{11}R_{12}R_{13}) \\ &\quad (R_{21} + R_{22} + R_{23} - R_{21}R_{22} - R_{21}R_{23} - R_{22}R_{23} + R_{21}R_{22}R_{23}) \\ &\quad (R_{31} + R_{32} + R_{33} - R_{31}R_{32} - R_{31}R_{33} - R_{32}R_{33} + R_{31}R_{32}R_{33}) \end{aligned}$$

When elements are identically distributed ($R_{im} \equiv R$), the reliability function R (R_1, \dots, R_9) is the component of SWCS and its structure function h for SWCS is given by

$$R(R_1, \dots, R_9) = 27R^3 - 81R^4 + 108R^5 - 81R^6 + 36R^7 - 9R^8 + R^9 \quad (42)$$

and

$$h(p_1, \dots, p_9) = 27p^3 - 81p^4 + 108p^5 - 81p^6 + 36p^7 - 9p^8 + p^9 \quad (43)$$

4.1 Signature of the SWCS

By using Owen's method for the components of SWCS, we get the structure function in the form of $h(y)$ as

$$h(y) = 27y^3 - 81y^4 + 108y^5 - 81y^6 + 36y^7 - 9y^8 + y^9 \quad (44)$$

Now using Eqs. (16) and (44), structure function can be expressed as

$$p(y) = y^9 h\left(\frac{1}{y}\right) = 1 - 9y + 36y^2 - 81y^3 + 108y^4 - 81y^5 + 27y^6 \quad (45)$$

Using Step 4 of Algorithm 2.2, we get tail signature \bar{B} for individual elements of the SWCS as

$$\bar{B}_1 = 1, \quad \bar{B}_2 = 1, \quad \bar{B}_3 = \frac{27}{28}, \quad \bar{B}_4 = \frac{6}{7}, \quad \bar{B}_5 = \frac{9}{14}, \quad \bar{B}_6 = \frac{9}{28}, \quad \bar{B}_7 = 0, \quad \bar{B}_8 = 0, \quad \bar{B}_9 = 0$$

Hence, tail signature of the SWCS is given by

$$\bar{B} = \left(1, 1, 1, \frac{27}{28}, \frac{6}{7}, \frac{9}{14}, \frac{9}{28}, 0, 0, 0 \right) \quad (46)$$

Now again using Step 5 of Algorithm 2.2, we obtain the signature of SWCS as

$$B = \left(0, 0, \frac{1}{28}, \frac{3}{28}, \frac{3}{14}, \frac{9}{28}, \frac{9}{28}, 0, 0 \right) \quad (47)$$

4.2 MTTF of the SWCS

MTTF is the meantime that causes system failure. It is also called failure rate of the component for system failure. The MTTF of reliability function $R(t)$ can be expressed as

$$R(t) = \int_0^\infty R_T(t) \quad (48)$$

The two methods namely, minimal signature and reliability function can be employed to calculate the expected lifetime based on the known parameters.

4.2.1 From expected lifetime

The expected lifetime of the SWCS for components can be evaluated using the Steps 2 and 3 of Algorithm 2.3.

By using Eq. (42), we get a minimal signature c as

$$c = (0, 0, 27, -81, 108, -81, 36, -9, 1)$$

of the SWCS elements.

The expected lifetime of SWCS of i.i.d. exponentially distributed elements with unit mean is obtained as

$$E(T) = 0.97$$

4.2.2 From reliability function

Reliability function $R(t)$ of the system can be evaluated using Step 4 of Algorithm 2.3 and then MTTF of the SWCS can be computed as

$$R(t) = 27R^3(t) - 81R^4(t) + 108R^5(t) - 81R^6(t) + 36R^7(t) - 9R^8(t) + R^9(t)$$

If elements are i.i.d. and exponentially distributed with an average one, then

$$R(t) = 27e^{-3t} - 81e^{-4t} + 108e^{-5t} - 81e^{-6t} + 36e^{-7t} - 9e^{-8t} + e^{-9t}$$

Therefore, MTTF of the system is obtained as

$$E(T) = 0.97 \quad (49)$$

It is worth mentioning that by both the methods we have got the exactly same value of $E(T)$.

4.3 Barlow-Proschan index

Now using the Algorithm 2.4, we can compute the Barlow-Proschan index ($I_{BP}^{(j)}$) from the reliability function with the help of Eq. (44) of the SWCS as

$$h(y) = 27y^3 - 81y^4 + 108y^5 - 81y^6 + 36y^7 - 9y^8 + y^9$$

Using Eq. (22), we get $I_{BP}^{(j)}$ for all the components of SWS as

$$I_{BP}^{(1)} = \int_0^1 (\partial_1 h)(y) dy = \int_0^1 (9y^2 - 36y^3 + 60y^4 - 54y^5 + 28y^6 - 8y^7 + y^8) dy = \frac{1}{9}$$

Similarly, we get Barlow-Proschan index $I_{BP}^l (l = 1, 2, \dots, 9)$ for individual components of the SWCS as

$$I_{BP} = \left(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9} \right)$$

4.4 Expected cost

We can evaluate the number of failed elements in the coherent system with irregular elements which cause system failure. Eryilmaz [22] discussed the expected value of the system signature using expected value of X . The average number of elements that are down at the moment of the system collapse can be determined by

$$E(T) = \sum_{j=1}^m c_j E_{X_{1:n}}$$

Using Step 1 of Algorithm 2.5, the expected value of X of the SWCS can be computed as

$$E(X) = \sum_{j=1}^9 j \cdot b_j, \quad j = 1, 2, \dots, 9$$

Hence, the expected value of X is obtained as

$$E(X) = 5.7857$$

Now, using Step 2 of Algorithm 2.5, we can compute the expected cost of SWCS as

$$\text{Expected cost} = \frac{E(X)}{E(T)} = 5.96465$$

5 Conclusion

In the present study, we have discussed the signature reliability characteristics of SWCS. The chapter aims to calculate the reliability function, signature, and minimal signature of SWS using U -function. In the considered model, SWCS has a subsystem which is connected in a parallel arrangement. This study also discusses the number of failed components when the system is failed by expected lifetime and Barlow-Proschan index method. The study reveals that this model is found to be more reliable on the basis of its signature reliability and expected lifetime. Further, it is observed that the Barlow-Proschan index is equal for every individual component and the expected lifetime calculated by two methods is equal.

References

- [1] G. Yingkui, L. Jing, Multi-state system reliability: a new and systematic review, *Procedia Eng.* 29 (2012) 531–536.
- [2] G. Levitin, Linear multi-state sliding-window systems, *IEEE Trans. Reliab.* 52 (2) (2003) 263–269.
- [3] A. Habib, R.O. Al-Seedy, T. Radwan, Reliability evaluation of multi-state consecutive k -out-of- r -from- n : G system, *Appl. Math. Model.* 31 (11) (2007) 2412–2423.
- [4] G. Levitin, Y. Dai, k -out-of- n Sliding window systems, *IEEE Trans. Syst. Man Cybern. Syst. Humans* 42 (3) (2012) 707–714.
- [5] H. Xiao, R. Peng, W. Wang, F. Zhao, Optimal element loading for linear sliding window systems, *Proc. Inst. Mech. Eng. Part O: J. Risk Reliab.* 230 (1) (2016) 75–84.
- [6] F.J. Samaniego, On closure of the IFR class under formation of coherent systems, *IEEE Trans. Reliab.* 34 (1) (1985) 69–72.
- [7] P.J. Boland, Signatures of indirect majority systems, *J. Appl. Probab.* 38 (2) (2001) 597–603.
- [8] G. Owen, Multilinear extensions and the Banzhaf value, *Naval Res. Logist. Quart.* 22 (4) (1975) 741–750.
- [9] G. Owen, Multilinear extensions of games, in: *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge, 1988, pp. 139–151.
- [10] R.E. Barlow, F. Proschan, Importance of system components and fault tree events, *Stochastic Process. Appl.* 3 (2) (1975) 153–173.
- [11] J. Navarro, J.M. Ruiz, C.J. Sandoval, Properties of coherent systems with dependent components, *Commun. Statist. Theory Meth.* 36 (1) (2007) 175–191.
- [12] J. Navarro, T. Rychlik, M. Shaked, Are the order statistics ordered? A survey of recent results, *Commun. Statist. Theory Meth.* 36 (7) (2007) 1273–1290.
- [13] F.J. Samaniego, *System Signatures and Their Applications in Engineering Reliability*, vol. 110, Springer Science & Business Media, New York, 2007.
- [14] X. Li, Z. Zhang, Some stochastic comparisons of conditional coherent systems, *Appl. Stoch. Model. Bus. Ind.* 24 (6) (2008) 541–549.
- [15] D. Bhattacharya, F.J. Samaniego, On the optimal allocation of components within coherent systems, *Statist. Probab. Lett.* 78 (7) (2008) 938–943.
- [16] I. Bairamov, B.C. Arnold, On the residual life lengths of the remaining components in an $n - k + 1$ out of n system, *Stat. Probab. Lett.* 78 (8) (2008) 945–952.
- [17] J. Navarro, R. Rubio, Computations of signatures of coherent systems with five components, *Commun. Statist. Simul. Comput.* 39 (1) (2009) 68–84.
- [18] J. Navarro, T. Rychlik, Comparisons and bounds for expected lifetimes of reliability systems, *Eur. J. Oper. Res.* 207 (1) (2010) 309–317.
- [19] J.L. Marichal, P. Mathonet, Computing system signatures through reliability functions, *Statist. Probab. Lett.* 83 (3) (2013) 710–717.
- [20] I. Ushakov, Universal generating function, *J. Comput. Sci. Syst.* 24 (1986) 118–129.
- [21] A. Høyland, M. Rausand, *System Reliability Theory: Models and Statistical Methods*, vol. 420, John Wiley & Sons, Hoboken, NJ, 2009.
- [22] S. Eryilmaz, The number of failed components in a coherent system with exchangeable components, *IEEE Trans. Reliab.* 61 (1) (2012) 203–207.

Series-parallel system study under warranty and preventive maintenance

5

N. Goyal, M. Ram

Graphic Era University, Dehradun, Uttarakhand, India

1 Introduction

The development of science and technology, and the essentials of modern society are competing against each other. That's why industries are trying to present more and more automation in their industrial developments, in order to fulfill the ever increasing requirements of society. The perfection in the effectiveness of series-parallel complex systems has a great significance due to various causes of failures like poor design, imperfect manufacturing, human error, under privileged maintenance, insufficient testing, inappropriate use, lack of environmental conditions, etc. The effectiveness of the system is its appropriateness for the fulfillment of the suggested tasks. The suitability of performing definite tasks is determined by the reliability [1–3]. To prevent the accidents and reduce the causes of failure, the system should be reliable and have a good quality. Reliability is basically a birth-to-death problem, including some zone like raw material and components, theoretical and comprehensive manufacturing design, fabrication, assessment and quality control, goods consignment in addition to packing, installation, operation, and preservation. Reliability of any system depends on various factors, most of which are random. The system can be made more reliable with enhanced manufacturing techniques [4,5].

Recently, the reliability of a series-parallel structure of the system with failed policy fascinates special attention because it has wide applications in every sphere of engineering sector such as industrial system, military system, communication system, and computer systems. It is one of the best ways of redundancy to improve both complex system reliability and the overall performance. Angus [6] modeled a k -out-of- n : G system into a parallel configuration with unlimited repair facility and derived an expression for the mean time between failure of the system with consideration of the exponential distribution of failures and repairs. Wu et al. [7] analyzed the reliability indices of the k -out-of- n : G structure of a repairable system with the consideration of single repair facility that can go to the single vacation and explained their study by taking the value of k as unity. Çekyay and Özekici [8] also studied the coherent system within the series configuration of the k -out-of- n structure of the standby systems and assumed that the component lifetimes are distributed exponentially. They derived the

generalized expressions of reliability measures. From this study, the performance investigation of the complex system with a large number of components is too much difficult.

Gupta and Gupta [9] developed an electronic system that consists of two subsystems in a series configuration including two types of failures namely unit and human, and one type of repair, and investigated the cost and economic benefits of the system through the availability and service cost. Kadyan and Ramniwas [10] developed a mathematical model of a single unit system using Markov process and supplementary variable technique and derived the expression for cost benefit using probabilistic indices including the warranty cost. Singh et al. [11] modeled a series system with two units in which each unit is controlled by its own controller and analyzed the cost benefit of the system including the service cost. Dhillon and Yang [12] presented a model that contains two identical units in parallel and one standby system with human errors and separates critical and noncritical human errors and a unit can fail either due to a nonhuman error or a human error. Singh et al. [13] modeled a repairable system that consists of two subsystems in a series configuration with their individual controller in which the first subsystem follows k -out-of- n : F redundancy and the second subsystem has a parallel configuration with two identical units by considering the two types of repair distribution and analyzed some reliability characteristics.

Ram and Singh [14] discussed a series-parallel complex repairable system, which consists of two independent subsystems A and B, which follow (1-out-of-2: F) and (1-out-of- n : F) redundancy, respectively, with two types of failure, partial and catastrophic and two types of repair, namely exponential and general and investigated the reliability indices.

In this work, authors designed a series-parallel system, in which subsystem A and subsystem B are in series configuration while subsystem A has two units, which are connected in parallel (Fig. 1). The designed system is analyzed for testing the performance and examined the reliability parameters and their sensitivity subjected to each failure rate of the system, expected profit, and warranty effect on it. The system is maintained with regular repair facility and preventive maintenance. A series-parallel system can be failed in many ways partially or completely. All the possible transitions of the designed system are shown in state transition diagram (Fig. 2). When the system is under preventive maintenance, the works stop completely, so it is considered as a complete failed state of the system.

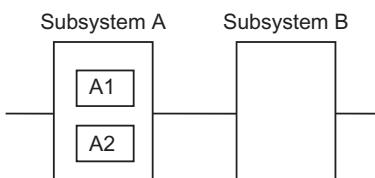


Fig. 1 Configuration diagram.

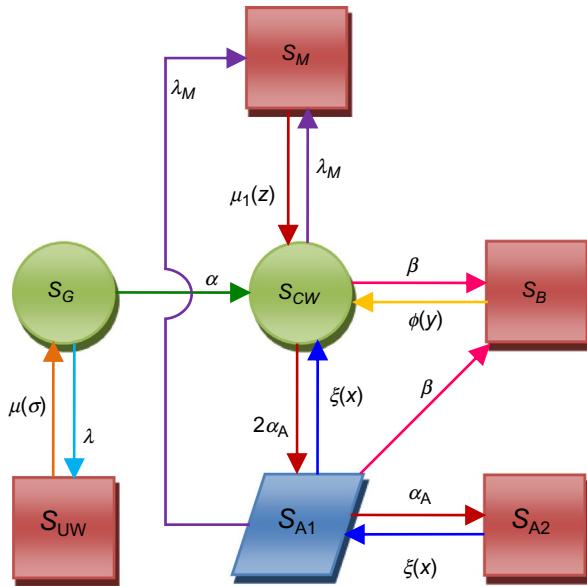


Fig. 2 Transition state diagram.

2 Mathematical modeling details

2.1 Assumptions and notations

The following assumptions have been used in the study of the proposed model and all the notations used throughout the work are shown in [Table 1](#).

- (i) Initially, the system is work free of failure.
- (ii) Only one change is allowed at a time in the transition states.
- (iii) All the components can be repaired.
- (iv) Sufficient repair facility is available.
- (v) When the failed component is repaired or after preventive maintenance, the system works like a new one.

2.2 State description

All the states of state transition diagram are described in [Table 2](#).

2.3 Formulation of the model

The following set of difference-differential equations by the probability of deliberation and steadiness of influences possesses the present mathematical model using Markov process.

Table 1 Notations

<i>t/s</i>	Time-scale/Laplace transform variable
$\bar{P}(s)$	Laplace transformation of $P(t)$
α	Rate of completion of warranty period
$\alpha_A/\beta/\lambda_M/\lambda$	Failure rates of subsystem A/subsystem B/system due to preventive maintenance/the system under warranty period
$P_i(t)$	The probability of the state S_i ; $i=0, 1, 2; j=G, CW, A1$
$P_j(\varepsilon, t)$	The probability density function of the state S_k at epoch t has an elapsed repair time of ε when $j=3, 4, 5, 6$ and $k=A2, B, M, UW$, respectively
$\xi(x)/\phi(y)/\mu_1(z)/\mu(\sigma)$	Repair rates for subsystem A/subsystem B/system due to preventive maintenance/the system under warranty period
w	Warranty period
$P_{up}(t)$	Up state system probability at time t or availability of the system
$Rl(t)$	The reliability of the system at time t
$E_p(t)$	Expected profit during the interval $[0,t)$
K_1/K_2	Revenue/service cost per unit time

Table 2 State description

S_G	Good working state of the system
S_{CW}	Good state of the system after completed warranty period
S_{A1}	Partially failed state of the system when one unit of subsystem A has been failed
S_{A2}	Complete failed state when subsystem A has been failed
S_B	Complete failed state when subsystem B has been failed
S_M	Complete failed state of the system under preventive maintenance
S_{UW}	Complete failed state of the system under warranty period

$$\left[\frac{\partial}{\partial t} + \alpha + \lambda \right] P_0(t) = \int_0^{\infty} \mu(\sigma) P_6(\sigma, t) d\sigma \quad (1)$$

$$\begin{aligned} \left[\frac{\partial}{\partial t} + 2\alpha_A + \beta + \lambda_M \right] P_1(t) = & \alpha P_0(t) + \xi(x) P_2(t) \\ & + \int_0^{\infty} \phi(y) P_4(y, t) dy + \int_0^{\infty} \mu_1(z) P_5(z, t) dz \end{aligned} \quad (2)$$

$$\left[\frac{\partial}{\partial t} + \alpha_A + \beta + \lambda_M + \xi(x) \right] P_2(t) = 2\alpha_A P_1(t) + \int_0^{\infty} \xi(x) P_3(x, t) dx \quad (3)$$

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \varepsilon} + A(\varepsilon) \right] P_i(\varepsilon, t) = 0; \quad i = 3, 4, 5, 6; \quad A = \xi, \phi, \mu_1, \mu; \quad \varepsilon = x, y, z, \sigma \quad (4)$$

Boundary conditions

$$P_3(0, t) = \alpha_A P_2(t) \quad (5)$$

$$P_4(0, t) = \beta [P_1(t) + P_2(t)] \quad (6)$$

$$P_5(0, t) = \lambda_M [P_1(t) + P_2(t)] \quad (7)$$

$$P_6(0, t) = \lambda P_0(t) \quad (8)$$

Initial condition $P_0(0) = 1$ and other state probabilities are zero at $t=0$.

2.4 Solution of the model

Taking the Laplace transformation of Eqs. (1)–(8) using initial condition

$$[s + \alpha + \lambda] \bar{P}_0(s) = 1 + \int_0^\infty \mu(\sigma) \bar{P}_6(\sigma, s) d\sigma \quad (9)$$

$$\begin{aligned} [s + 2\alpha_A + \beta + \lambda_M] \bar{P}_1(s) &= \alpha \bar{P}_0(s) + \xi(x) \bar{P}_2(s) \\ &\quad + \int_0^\infty \phi(y) \bar{P}_4(y, s) dy + \int_0^\infty \mu_1(z) \bar{P}_5(z, s) dz \end{aligned} \quad (10)$$

$$[s + \alpha_A + \beta + \lambda_M + \xi(x)] \bar{P}_2(s) = 2\alpha_A \bar{P}_1(s) + \int_0^\infty \xi(x) \bar{P}_3(x, s) dx \quad (11)$$

$$\left[s + \frac{\partial}{\partial \varepsilon} + A(\varepsilon) \right] \bar{P}_i(\varepsilon, s) = 0; \quad i = 3, 4, 5, 6; \quad A = \xi, \phi, \mu_1, \mu; \quad \varepsilon = x, y, z, \sigma \quad (12)$$

$$\bar{P}_3(0, s) = \alpha_A \bar{P}_2(s) \quad (13)$$

$$\bar{P}_4(0, s) = \beta [\bar{P}_1(s) + \bar{P}_2(s)] \quad (14)$$

$$\bar{P}_5(0, s) = \lambda_M [\bar{P}_1(s) + \bar{P}_2(s)] \quad (15)$$

$$\bar{P}_6(0, s) = \lambda \bar{P}_0(s) \quad (16)$$

After solving Eqs. (9)–(16), the expression for each transition-state probabilities is derived as

$$\bar{P}_0(s) = \frac{1}{s + \alpha + \lambda (1 - \bar{S}_\mu(s))} \quad (17)$$

$$\bar{P}_1(s) = \frac{\alpha}{B_3} \bar{P}_0(s) \quad (18)$$

$$\bar{P}_2(s) = \frac{\alpha}{B_3} \left(\frac{2\alpha_A}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right) \bar{P}_0(s) \quad (19)$$

$$\bar{P}_3(s) = \left(\frac{1 - \bar{S}_\xi(s)}{s} \right) \frac{2(\alpha_A)^2 \alpha}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \bar{P}_0(s) \quad (20)$$

$$\bar{P}_4(s) = \left(\frac{1 - \bar{S}_\phi(s)}{s} \right) \left(1 + \frac{2\alpha_A}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right) \frac{\alpha\beta}{B_3} \bar{P}_0(s) \quad (21)$$

$$\bar{P}_5(s) = \left(\frac{1 - \bar{S}_{\mu_1}(s)}{s} \right) \left(1 + \frac{2\alpha_A}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right) \frac{\alpha\lambda_M}{B_3} \bar{P}_0(s) \quad (22)$$

$$\bar{P}_6(s) = \left(\frac{1 - \bar{S}_\mu(s)}{s} \right) \lambda \bar{P}_0(s) \quad (23)$$

With the help of all transition-state probabilities, the expression of the Laplace transformation of the probability of upstate and downstate system has been obtained as follows:

$$\bar{P}_{up}(s) = \sum_{i=0}^2 \bar{P}_i(s) = \left[1 + \frac{\alpha}{B_3} \left(1 + \frac{2\alpha_A}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right) \right] \bar{P}_0(s) \quad (24)$$

$$\begin{aligned} \bar{P}_{down}(s) &= \sum_{i=3}^6 \bar{P}_i(s) \\ &= \left[\frac{\alpha}{B_3} \left(1 + \frac{2\alpha_A}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right) \left\{ \alpha \left(\frac{1 - \bar{S}_\phi(s)}{s} \right) + \lambda_M \left(\frac{1 - \bar{S}_{\mu_1}(s)}{s} \right) \right\} \right. \\ &\quad \left. + \left(\frac{1 - \bar{S}_\mu(s)}{s} \right) \lambda + \left(\frac{1 - \bar{S}_\xi(s)}{s} \right) \frac{2(\alpha_A)^2 \alpha}{s+B_2 - \alpha_A \bar{S}_\xi(s)} \right] \bar{P}_0(s) \end{aligned} \quad (25)$$

3 Particular cases and numerical computations

3.1 Availability analysis

Availability of a system can be developed by an affected plan that focuses on growing testability and maintainability and not on reliability. The availability of the system depends on the system organization as well as the component's availability [2]. Setting the failure and repair rates as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$,

$\lambda = 0.02$, $\mu(\sigma) = 1$, $\mu_1(z) = 1$, $\phi(y) = 1$, $\xi(x) = 1$ [10,11,15] in Eq. (24) and by taking the inverse Laplace transformation, one can obtain the availability of the considered series-parallel system as

$$\begin{aligned} P_{up}(t) = & 0.08633 \cosh(0.47170)e^{(-0.55000t)} + 0.00088e^{(-1.35324t)} \\ & - 0.00913e^{(-1.08696)} + 0.92231 - 0.00038e^{(-0.86979t)} \\ & + 0.03903e^{(-0.55000t)} \sinh(0.47170) \end{aligned} \quad (26)$$

The availability of the proposed system is shown in [Table 3](#) and the corresponding [Fig. 3](#), after varying the time unit t from 0 to 60 with difference 5, respectively.

3.2 Reliability analysis

System reliability is one of the key elements which is in agreement with the performance of each component of the system. Reliability represents the probability of non-failure components, subsystems, and system to perform their required functions for a precise time period in specified environmental condition [12,16]. Repair facility has not taken into account the reliability evaluation, so, setting all repair rates as zero and assigning the value of failure rates as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$, $\lambda = 0.02$ [10,11,15] in Eq. (24), later on taking the inverse Laplace transformation, one can get the reliability expression in terms of time as

$$R(t) = e^{(-0.18000t)} - 5.33333e^{(-0.13000t)} + 5.33333e^{(-0.10000t)} \quad (27)$$

By varying the time unit t from 0 to 60, one can determine the reliability of the considered system as revealed in [Table 4](#) and corresponding [Fig. 4](#).

Table 3 Availability as function of time

Time (t)	Availability $P_{up}(t)$
0	1.00000
5	0.96478
10	0.95095
15	0.94167
20	0.93540
25	0.93116
30	0.92829
35	0.92635
40	0.92504
45	0.92416
50	0.92356
55	0.92315
60	0.92288

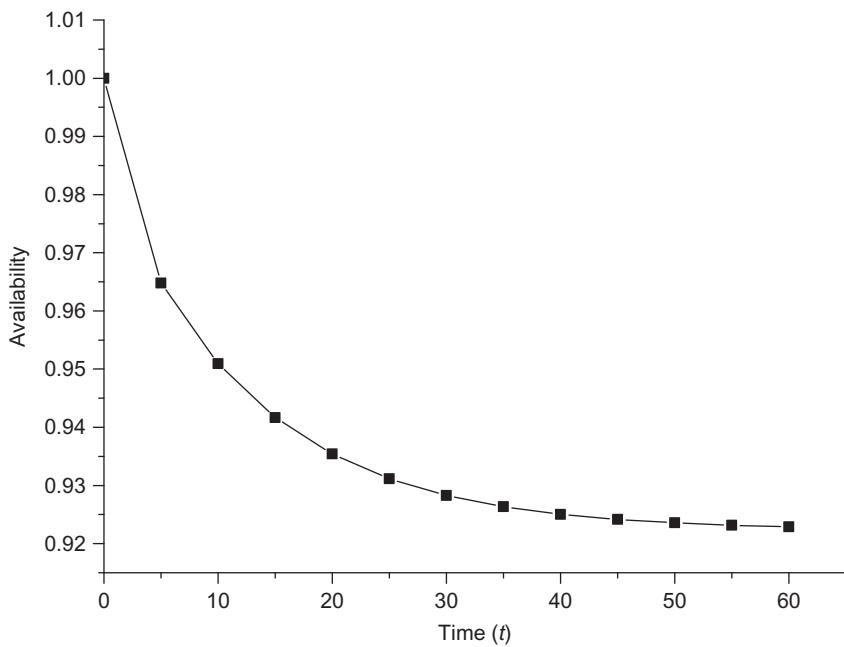


Fig. 3 Availability as function of time.

Table 4 Reliability as function of time

Time (t)	Reliability $Rl(t)$
0	1.00000
5	0.85715
10	0.67382
15	0.49844
20	0.35299
25	0.24210
30	0.16209
35	0.10653
40	0.06901
45	0.04419
50	0.02804
55	0.01766
60	0.01105

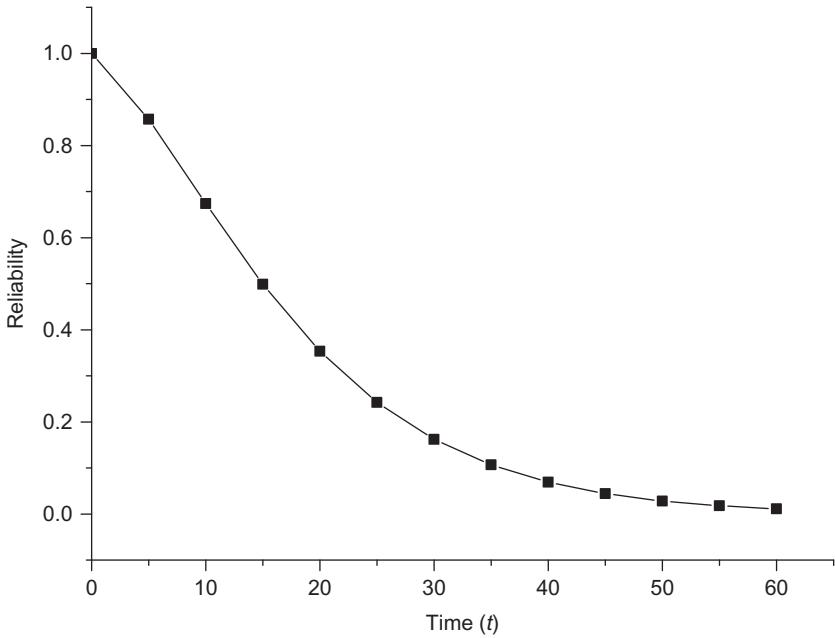


Fig. 4 Reliability as function of time.

3.3 MTTF (Mean Time to Failure) analysis

By considering the repair rate zero and taking the limits of s tends to zero in Eq. (24), one can obtain the MTTF of the system as

$$\begin{aligned} \text{MTTF} &= \lim_{s \rightarrow 0} \bar{P}_{up}(s) \\ &= \frac{1}{\alpha + \lambda} \left[1 + \frac{\alpha}{2\alpha_A + \beta + \lambda_M} \left(1 + \frac{2\alpha_A}{\alpha_A + \beta + \lambda_M} \right) \right] \end{aligned} \quad (28)$$

Now varying the failure rates one by one at 0.01, 0.02, ..., 0.09, respectively and setting the other failure rates as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$, $\lambda = 0.02$ [17,18] in Eq. (28), MTTF with the variation in failure rates is signified in Table 5 and their graphical representation are presented in Fig. 5.

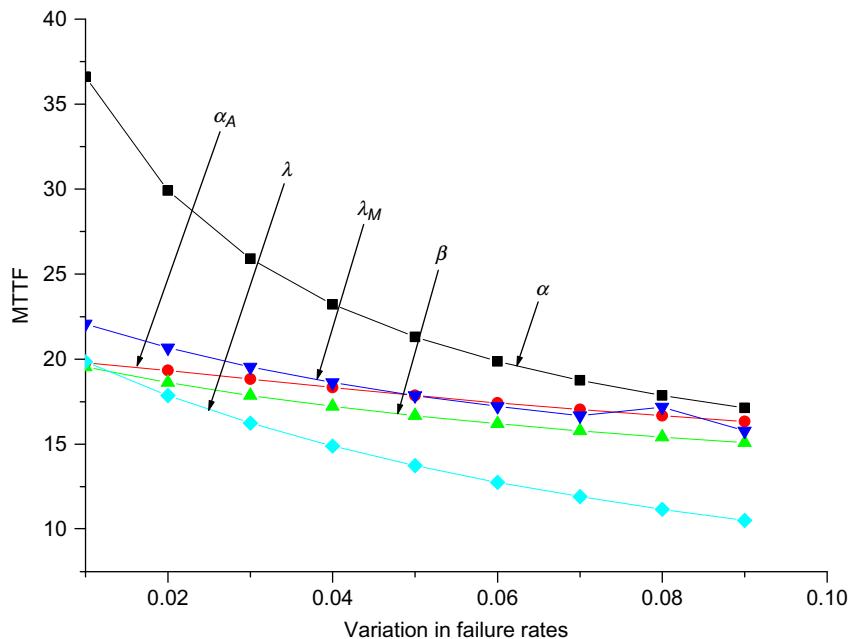
3.4 Expected profit

Let the service facility be always available; the expected profit during the interval $[0, t]$ is given as:

$$E_P(t) = K_1 \int_0^t P_{up}(t) dt - (t - w) K_2 \quad (29)$$

Table 5 MTTF as function of failure rates

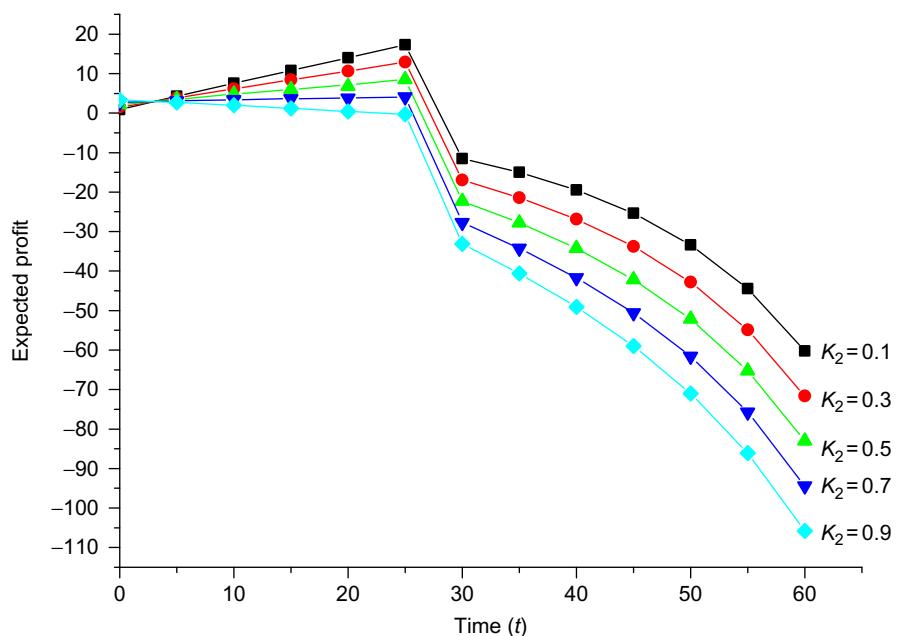
Variation in $\alpha, \alpha_A, \beta, \lambda_M, \lambda$	MTTF with respect to failure rates				
	α	α_A	β	λ_M	λ
0.01	36.60969	19.77778	19.54545	22.06349	19.84805
0.02	29.91453	19.33334	18.62745	20.66666	17.86325
0.03	25.89743	18.83117	17.86325	19.54545	16.23932
0.04	23.21937	18.33334	17.21804	18.62745	14.88604
0.05	21.30647	17.86325	16.66667	17.86325	13.74096
0.06	19.87179	17.42857	16.19048	17.21804	12.75946
0.07	18.75593	17.03030	15.77540	16.66667	11.90883
0.08	17.86325	16.66667	15.41063	17.19047	11.16453
0.09	17.13287	16.33484	15.08772	15.77540	10.50779

**Fig. 5** MTTF as function of failure rates.

Using Eq. (26) in Eq. (29) and integrating with respect to t , one can obtain the expected profit of the designed system as shown in Table 6 and corresponding Fig. 6, after considering the value of $K_1 = 1$, warranty period $[0, 3]$ and the value of K_2 varies as 0.1, 0.3, 0.5, 0.7, 0.9, respectively, in Eq. (29), [10,11].

Table 6 Expected profit as function of time

Time (t)	Expected profit				
	$K_2 = 0.1$	$K_2 = 0.3$	$K_2 = 0.5$	$K_2 = 0.7$	$K_2 = 0.9$
0	0.90000	1.50000	2.10000	2.70000	3.30000
5	4.28572	3.88572	3.48572	3.08572	2.68572
10	7.57267	6.17267	4.77267	3.37267	1.97267
15	10.80273	8.40273	6.00273	3.60273	1.20273
20	13.99465	10.59465	7.19465	3.79465	0.39465
25	17.32449	12.92449	8.52449	4.12449	-0.27551
30	-11.51548	-16.91548	-22.31548	-27.71548	-33.11548
35	-15.01227	-21.41227	-27.81227	-34.21227	-40.61227
40	-19.47588	-26.87588	-34.27588	-41.67588	-49.07588
45	-25.36639	-33.76639	-42.16639	-50.56639	-58.96639
50	-33.36539	-42.76539	-52.16539	-61.56539	-70.96539
55	-44.48181	-54.88181	-65.28181	-75.68181	-86.08181
60	-60.20855	-71.60855	-83.00855	-94.40855	-105.80855

**Fig. 6** Expected profit versus time.

3.5 Sensitivity analysis

Sensitivity analysis is a measure that finds out how the uncertainty in the output or sensitive an output of a mathematical model or a complex system can be assigned to different sources of uncertainty in its inputs or to change in an input while keeping other inputs constant [19,20]. The sensitivity of a measure is defined as the partial derivative of the function with respect to their input factors. Here, these input factors are failure rates of the designed system.

3.5.1 Availability sensitivity

One can determine the availability sensitivity of the designed system in concern of each failure rate by taking the inverse Laplace transform of Eq. (24) and differentiating partially with respect to the input parameters (failure rates). Substituting the value of failure and repair rates as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$, $\lambda = 0.02$, $\mu(\sigma) = 1$, $\mu_1(z) = 1$, $\phi(y) = 1$, $\xi(x) = 1$, in that partial derivatives. Table 7 and Fig. 7 show the availability sensitivity of the designed system at a different instance.

3.5.2 Reliability sensitivity

Reliability sensitivity can be obtained by differentiating partially Eq. (24) with respect to their input parameters; after taking the inverse Laplace transformation and fixing the values of input parameters as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$, $\lambda = 0.02$, $\mu(\sigma) = 0$, $\mu_1(z) = 0$, $\phi(y) = 0$, $\xi(x) = 0$ in the partial derivatives, one can obtain Table 8 and correspondingly Fig. 8.

Table 7 Availability sensitivity as function of time

Time (t)	Availability sensitivity				
	$\frac{\partial P_{up}(t)}{\partial \alpha}$	$\frac{\partial P_{up}(t)}{\partial \alpha_A}$	$\frac{\partial P_{up}(t)}{\partial \beta}$	$\frac{\partial P_{up}(t)}{\partial \lambda_M}$	$\frac{\partial P_{up}(t)}{\partial \lambda}$
0	0.00000	0.00000	0.00000	0.00000	0.00000
5	-0.16555	-0.03128	-0.23589	-0.23589	-0.68525
10	-0.25167	-0.06569	-0.43404	-0.43404	-0.45527
15	-0.26474	-0.08914	-0.56802	-0.56803	-0.29976
20	-0.24293	-0.10499	-0.65859	-0.65860	-0.19722
25	-0.20747	-0.11571	-0.71982	-0.71983	-0.12966
30	-0.16948	-0.12295	-0.76122	-0.76122	-0.08517
35	-0.13434	-0.12785	-0.78920	-0.78920	-0.05590
40	-0.10418	-0.13116	-0.80812	-0.80812	-0.03666
45	-0.079458	-0.13340	-0.82091	-0.82091	-0.02401
50	-0.059823	-0.13491	-0.82955	-0.82956	-0.01572
55	-0.044569	-0.13593	-0.83540	-0.83540	-0.01027
60	-0.032921	-0.13662	-0.83935	-0.83935	-0.06708

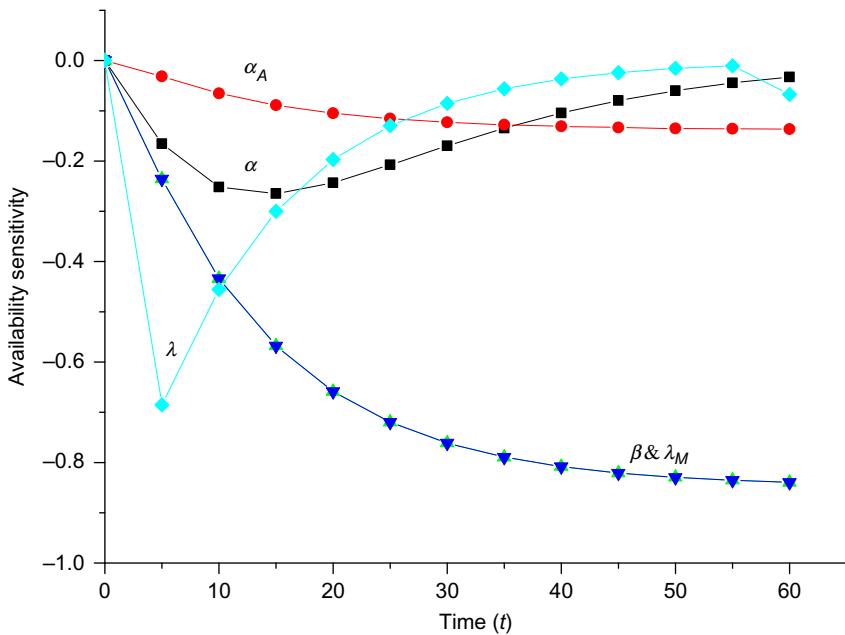


Fig. 7 Availability sensitivity as function of time.

Table 8 Reliability sensitivity as function of time

Time (t)	Reliability sensitivity				
	$\frac{\partial RI(t)}{\partial \alpha}$	$\frac{\partial RI(t)}{\partial \alpha_A}$	$\frac{\partial RI(t)}{\partial \beta}$	$\frac{\partial RI(t)}{\partial \lambda_M}$	$\frac{\partial RI(t)}{\partial \lambda}$
0	0.00000	0.00000	0.00000	0.00000	0.00000
5	-0.52131	-0.16498	-0.63165	-0.63165	-3.65413
10	-1.37755	-0.65713	-1.53640	-1.53640	-5.20180
15	-1.98371	-1.11054	-2.05151	-2.05151	-5.42505
20	-2.20986	-1.32556	-2.12923	-2.12923	-4.93049
25	-2.13198	-1.31085	-1.92033	-1.92033	-4.13217
30	-1.87554	-1.15300	-1.58338	-1.58338	-3.27934
35	-1.54738	-0.93678	-1.22702	-1.22702	-2.50155
40	-1.21790	-0.71902	-0.90877	-0.90877	-1.85156
45	-0.92474	-0.52896	-0.65036	-0.65036	-1.33828
50	-0.68260	-0.37664	-0.45315	-0.45315	-0.94889
55	-0.49257	-0.26138	-0.30908	-0.30908	-0.66224
60	-0.34890	-0.17770	-0.20720	-0.20720	-0.45610

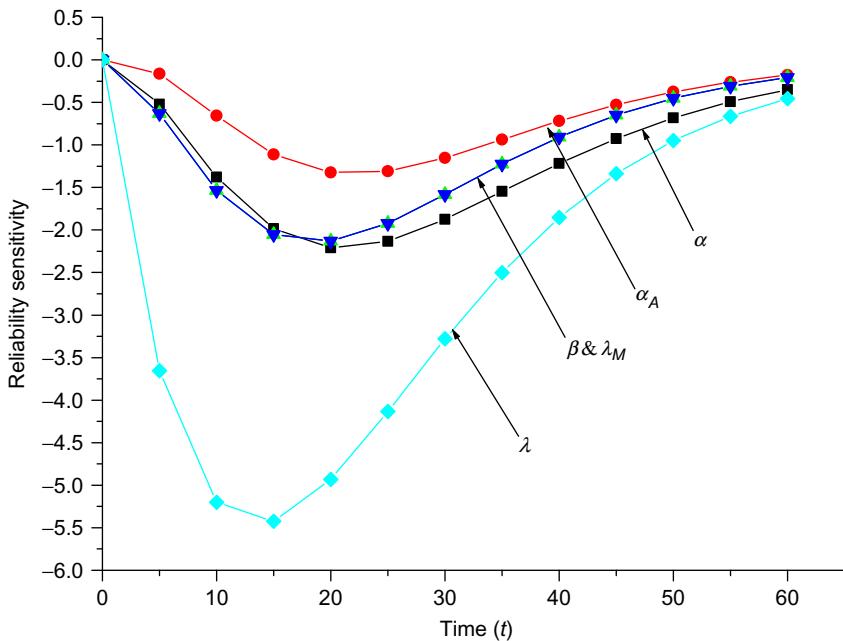


Fig. 8 Reliability sensitivity as function of time.

3.5.3 MTTF sensitivity

MTTF sensitivity can be computed by partial differentiation of Eq. (28) with respect to the input parameters and then varying input parameters one by one at 0.01, 0.02, ..., 0.09, respectively, and setting the other failure rates as $\alpha = 0.08$, $\beta = 0.03$, $\lambda_M = 0.05$, $\alpha_A = 0.05$, $\lambda = 0.02$ in partial derivatives of Eq.(28), one may get Table 9 and corresponding Fig. 9.

Table 9 MTTF sensitivity as function of failure rates

Variation in λ_{HE} , α_A , β , λ_M , λ	MTTF sensitivity				
	$\frac{\partial \text{MTTF}}{\partial \alpha}$	$\frac{\partial \text{MTTF}}{\partial \alpha_A}$	$\frac{\partial \text{MTTF}}{\partial \beta}$	$\frac{\partial \text{MTTF}}{\partial \lambda_T}$	$\frac{\partial \text{MTTF}}{\partial \lambda}$
0.01	-892.68756	-37.53086	-100.98140	-156.71453	-220.53392
0.02	-502.13675	-48.88889	-83.42945	-124.44444	-178.63248
0.03	-321.36752	-50.59875	-69.98320	-100.98140	-147.63015
0.04	-223.17189	-48.61111	-59.47199	-83.42945	-124.05033
0.05	-163.96302	-45.29184	-51.11111	-69.98320	-105.69969
0.06	-125.53419	-41.63265	-44.35941	-59.47199	-91.13902
0.07	-99.18751	-38.05326	-38.83440	-51.11111	-79.39221
0.08	-80.34188	-34.72222	-34.25984	-44.35941	-69.77831
0.09	-66.39825	-31.69468	-30.43244	-38.83440	-61.81055

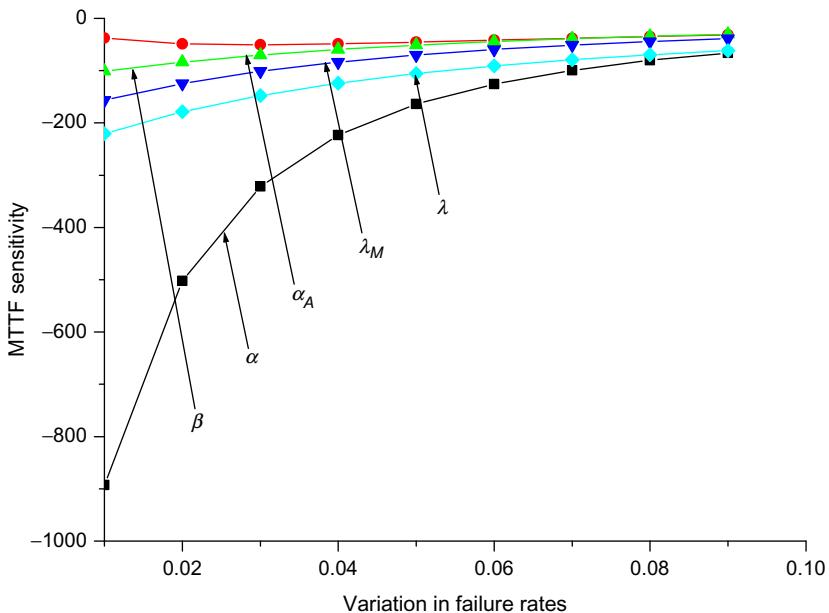


Fig. 9 MTTF sensitivity as function of failure rates.

4 Results discussion

In this research work, authors have studied the various reliability characteristics such as availability, reliability, MTTF, expected profit and sensitivity of availability, reliability, and MTTF of the system under the consideration of distinct failures. Through the overall study on the series-parallel system, authors made the following analysis.

Figs. 3 and 4 give the idea about the availability and reliability of the system, respectively. The graphs of availability and reliability show that the availability and reliability of the proposed system decrease smoothly as time increases. After a long time period, the availability of the system will establish at a fixed point while its reliability tends to zero.

The trend of mean time to failure of the system in concern of each failure rate is demonstrated by the graphs as shown in Fig. 5. From the study of these graphs, authors analyze that the MTTF of the system decreases with respect to each failure rate as failure rate increases. For the series-parallel system, MTTF is lowest with respect to the failure rate of the system under warranty while it is highest in the case of the rate of successful completion of warranty.

The profit in the system is affected by the various costs such as regular maintenance cost, warranty cost, and preventive maintenance cost. The investigation of the cost benefit is revealed graphically in Fig. 6 and gives that the profit of the system decreases as service cost increases while it increases as time increases. From the critical examination of this graph, authors analyze that after approximately 25 months, the system goes into the loss.

The behavior of sensitivity analysis of availability, reliability, and MTTF of the designed system is revealed in [Figs.7–9](#), respectively. The vital examination of [Fig. 9](#) epitomizes that the availability sensitivity of the system first decreases swiftly with respect to failure rate of the system under warranty and after 5 months it increases smoothly as time increases while the sensitivity of the availability with respect to the rate of completion of warranty period first decreases smoothly and after a short period it increases in a similar manner as time increases. The sensitivity availability of the system decreases with respect to the other failure rates as time increases. Similarly, from the graph of reliability sensitivity, it can be seen that the sensitivity in the reliability of the system first decreases smoothly and after a short period it increases in the same discipline with respect to each failure rate as time increase. With the precarious study of [Fig. 9](#), one can see that the sensitivity of MTTF increases slightly with respect to the failure rate of the system under warranty, subsystem B, and preventive maintenance. In the situation of variation in the failure rate of subsystem A, the sensitivity of MTTF decreases slightly while the rate of completion of warranty period increases, it increases smoothly.

5 Conclusion

The mathematical model of a series-parallel system has been designed in this research work to determine the reliability measures, which help to analyze the system performance. The proposed system is analyzed under the consideration of equipment failure and preventive maintenance with the fixed warranty. After the investigation, authors concluded that the reliability and MTTF of the system can be improved by controlling the system failures while availability can be enhanced with the better maintenance facility. It is also concluded that one can attain maximum profit with a long time warranty and by taking care of service cost. The series-parallel system is most sensitive in case of the parallel subsystem i.e. subsystem A and least sensitive in case of failure under warranty. It is worthwhile that this study is very beneficial for the designers, engineers, and system operators to attain the highly reliable system. In future, authors will try to investigate some other factors such as mean time to repair and mean time between maintenance of the series-parallel system.

References

- [1] M. Ram, Reliability measures of a three-state complex system: a copula approach, *Appl. Appl. Math. Int. J.* 5 (10) (2010) 1483–1492.
- [2] A. Avizienis, J.C. Laprie, B. Randell, C. Landwehr, Basic concepts and taxonomy of dependable and secure computing, *IEEE Trans. Depend. Secure Comput.* 1 (1) (2004) 11–33.
- [3] B.S. Dhillon, C. Singh, *Engineering Reliability: New Techniques and Applications*, Wiley, New York, 1981. pp. 329–334.
- [4] B.S. Dhillon, *Reliability, Quality, and Safety for Engineers*, CRC Press, Boca Raton, FL, 2004.

- [5] M. Ram, N. Goyal, Gas turbine assimilation under copula-coverage approaches, in: Research Advances in Industrial Engineering, Springer International Publishing, Switzerland, 2015, , pp. 103–116.
- [6] J.E. Angus, On computing MTBF for a k -out-of- n : G repairable system, IEEE Trans. Reliab. 37 (3) (1988) 312–313.
- [7] W. Wu, Y. Tang, M. Yu, Y. Jiang, Reliability analysis of a k -out-of- n : G repairable system with single vacation, Appl. Math. Model. 38 (24) (2014) 6075–6097.
- [8] B. Çekyay, S. Özekici, Reliability, MTTF and steady-state availability analysis of systems with exponential lifetimes, Appl. Math. Model. 39 (1) (2015) 284–296.
- [9] P.P. Gupta, R.K. Gupta, Cost analysis of an electronic repairable redundant system with critical human errors, Microelectron. Reliab. 26 (3) (1986) 417–421.
- [10] M.S. Kadyan, Ramniwas, Cost benefit analysis of a single-unit system with warranty for repair, Appl. Math. Comput. 223 (2013) 346–353.
- [11] V.V. Singh, M. Ram, D.K. Rawal, Cost analysis of an engineering system involving subsystems in series configuration, IEEE Trans. Autom. Sci. Eng. 10 (4) (2013) 1124–1130.
- [12] B.S. Dhillon, N. Yang, Availability of a man-machine system with critical and non-critical human error, Microelectron. Reliab. 33 (1993) 1511–1521.
- [13] V.V. Singh, S.B. Singh, M. Ram, C.K. Goel, Availability, MTTF and cost analysis of a system having two units in series configuration with controller, Int. J. Syst. Assur. Eng. Manag. (2012) 1–12.
- [14] M. Ram, S.B. Singh, Availability, MTTF and cost analysis of complex system under preemptive-repeat repair discipline using Gumbel-Hougaard family copula, Int. J. Qual. Reliab. Manag. 27 (5) (2010) 576–595.
- [15] M. Ram, S.B. Singh, V.V. Singh, Stochastic analysis of a standby system with waiting repair strategy, IEEE Trans. Syst. Man Cybern. Syst. 43 (3) (2013) 698–707.
- [16] W.Q. Meeker, L.A. Escobar, Reliability: the other dimension of quality, Qual. Technol. Quant. Manag. 1 (1) (2003) 1–25.
- [17] M. Manglik, M. Ram, Behavioural analysis of a hydroelectric production power plant under reworking scheme, Int. J. Prod. Res. 53 (2) (2015) 648–664.
- [18] N. Goyal, M. Ram, A. Kaushik, Automotive water cooling system analysis subject to time dependence and failure issues, Int. J. Manuf. Mater. Mech. Eng. 6 (2) (2016) 22.
- [19] A. Saltelli, K. Chan, E.M. Scott, Sensitivity Analysis, vol. 1, Wiley, New York, 2000.
- [20] U.G.U.S.T.O. Gandini, Importance and sensitivity analysis in assessing system reliability, IEEE Trans. Reliab. 39 (1) (1990) 61–70.

On solving complex reliability optimization problem using multi-objective particle swarm optimization

6

A. Kumar*, S. Pant*, M. Ram[†], S.B. Singh[‡]

*University of Petroleum & Energy Studies, Dehradun, India, [†]Graphic Era University, Dehradun, India, [‡]G. B. Pant University of Agriculture and Technology, Pantnagar, India

1 Introduction

Reliability is always a top customer concern and is increasingly vocalized by customers as a major factor in purchasing decisions. When someone assigns attribute “reliable” to a component or a system, it precisely means to say that the same will render service for a good or at least reasonable period of time [1]. Designing a highly reliable system is the most challenging task for a design engineer because it is mandatory for him/her to strike a balance between multiple competing objectives like to maximize reliability, minimize cost, minimize weight, etc. Optimization deals with finding the extreme (optimal) value of a function in a domain of definition, subject to various constraints on the variable values. In other words, optimization is the act of determining the value of certain parameters subject to constraints so that some measure of optimality is satisfied. The applicability of optimization methods is widespread (Science, Engineering, Mathematics, Economics, Commerce, Management etc.), reaching into almost every activity in which numerical information is processed. In single objective optimization problem (SOOP), the main objective is to find one optimum solution. Therefore, it cannot provide a set of alternative solutions that trade different objective against each other. On the contrary, the main objective of a multi-objective optimization problem (MOOP) is to find a set of compromised solutions (Pareto-optimal solutions) instead of finding a single optimal solution. These Pareto-optimal solutions help DM to find the most preferred optimal solution according to his/her subjective preferences [2]. As far as solution to multi-objective problem is concerned one can solve such problems either under single objective formulation or under multi-objective formulation. In the past two decades the problems of multi-objective reliability optimization have been extensively solved [3–12].

In general, obtaining optimal reliability design is a tedious task because of the non-deterministic polynomial-time hard (NP-hard) nature of reliability optimization problems [13]. So it is almost impossible to find the solution of those types of problems using exact methods or heuristics. Therefore, metaheuristic algorithms, particularly,

particle swarm optimization (PSO), gray wolf optimization (GWO) algorithm, genetic algorithm (GA), Cuckoo Search algorithm (CSA), ant colony optimization (ACO), etc. are suitable for solving reliability optimization problems. Recently, many meta-heuristics [2,14,15] have been employed to solve reliability optimization problems.

1.1 Multi-objective reliability optimization problems

$$\text{Max } F = (f_1(r_1, r_2, \dots, r_n, x_1, x_2, \dots, x_n), f_2(r_1, r_2, \dots, r_n, x_1, x_2, \dots, x_n), \dots, f_K(r_1, r_2, \dots, r_n, x_1, x_2, \dots, x_n))$$

subject to

$$f_i^c(r_1, r_2, \dots, r_n, x_1, x_2, x_n) \leq b_i, \quad \text{for } i = 1, 2, \dots, m$$

$$l_j \leq x_j \leq u_j, x_j \in Z^+, \quad \text{for } j = 1, 2, \dots, n$$

$$r_j \in (0, 1) \subset R, \quad \text{for } j = 1, 2, \dots, n$$

where $f_k, \forall k = 1, 2, \dots, K$ is one of the objective functions of the problem and K is the total number of objective functions. In most practical situations involving reliability optimization, there are several mutually conflicting goals such as maximizing system reliability and minimizing cost, weight, volume, and constraints required to be addressed simultaneously. Some main objectives can be expressed as:

Objective 1 The most important objective is the maximization of system reliability (R_S). It enables the system to function satisfactorily throughout its intended service period

$$\text{Max } R_S$$

As in our approach we are considering all minimization problems. Hence, the above objective is equivalent to minimization of system unreliability ($Q_S = 1 - R_S$) that can be expressed as follows

$$\text{Min } Q_S$$

Objective 2 The addition of the redundant components increases not only the system reliability but also its overall cost (C_S). A manufacturer has to balance these conflicting objectives, keeping in view the importance of reducing the overall cost

$$\text{Min } C_S$$

Objective 3 As with cost, every added redundant component increases the weight of the system. Usually, the overall weight of a system needs to be minimized along with its cost even as reliability is maximized (or unreliability is minimized)

$$\text{Min } W_S$$

1.2 Pareto dominance

For any minimization problem, a solution \vec{x}^1 is said to dominate \vec{x}^2 (denoted by $\vec{x}^1 \prec \vec{x}^2$) iff \vec{x}^1 is no worse than \vec{x}^2 in all objectives, i.e. $f_i(\vec{x}^1) \leq f_i(\vec{x}^2), \forall i = 1, 2, 3, \dots, M$, and \vec{x}^1 is strictly better than \vec{x}^2 in at least one objective, i.e. $\exists i = 1, 2, 3, \dots, M : f_i(\vec{x}^1) < f_i(\vec{x}^2)$.

The definition for a maximization problem (\succ) is analogical. If \vec{x}^1 dominates \vec{x}^2 , it is also customary to write any of the following:

- \vec{x}^2 is dominated by \vec{x}^1 ;
- \vec{x}^1 is nondominated by \vec{x}^2 , or \vec{x}^1 is noninferior to \vec{x}^2 .

1.3 Particle swarm optimization

PSO, developed by Eberhart and Kennedy, is a swarm intelligence method for solving optimization problems [1]. PSO is inspired by the ability of flocks of birds, schools of fish, and herds of animals to adapt to their environment, find rich sources of food, and avoid predators by implementing “information sharing” approaches, hence, developing an evolutionary advantage [11,16,17]. PSO is initialized with randomly generated population of particles (initial swarm) and a random velocity is assigned to each particle that propagates the particle in search space toward optima over a number of iterations. Each particle has a memory remembering best position attained by it in the past, which is called personal best position (P_{best}). Each particle has its P_{best} and the particle with the best value of fitness is called global best particle (G_{best}). Suppose that the search space is D dimensional, the i th particle of the population can be represented by a D -dimensional vector $(x_i^1, x_i^2, \dots, x_i^D)^T$. The velocity of this particle can be represented by another D -dimensional vector $(V_i^1, V_i^2, \dots, V_i^D)^T$. The previously best visited position of i th particle is denoted by P_i and the best particle in the swarm is denoted by P_g . The update of the particle’s position is accomplished by the following two equations. Eq. (1) calculates a new velocity for each particle based on its previous velocity and Eq. (2) updates each particle’s position in search space.

$$V_{id}^{k+1} = wV_{id}^k + c_1 r_1 [p_{id}^k(t) - x_{id}(t)] + c_2 r_2 [p_g^k(t) - x_{id}^k(t)] \quad (1)$$

$$x_{id}^{k+1}(t+1) = x_{id}^k(t) + v_{id}^{k+1}(t+1) \quad (2)$$

where k =iteration number, $d=1,2,3, \dots, D$; $i=1,2,3, \dots, N$; N =swarm size, w =inertia weight, which controls the momentum of particle by weighing the contribution of previous velocity; c_1 and c_2 are positive constants called acceleration coefficients; r_1 and r_2 are random numbers uniformly distributed between [0,1].

2 Multi-objective particle swarm optimization incorporating the mechanism of crowding distance

MOPSO-CD is a swarm-based artificial intelligence-based technique. It is inspired from the food searching behavior of birds flocking or fish schooling. Crowding distance technique has been extensively applied in evolutionary multi-objective optimization algorithms to promote the diversity. The proposal for using crowding distance measure in MOPSO for G_{best} selection and archiving updating was first made by Raquel et al. [6]. This approach used a mutation operator proposed by Coello et al. [18] to produce a highly explorative behavior in the algorithm in such a way that it is applied only during a certain number of generations at the beginning of the process. This is helpful in terms of preventing premature convergence due to existing local Pareto optimal fronts in some optimization problems. The performance of this approach was highly competitive in converging toward the Pareto optimal front and generated a well-distributed set of nondominated solutions. MOPSO-CD has drawn some attention recently as it exhibits a relatively fast convergence and well-distributed Pareto optimal front compared with other multi-objective optimization algorithms [19,20]. Therefore, MOPSO-CD approach has been applied to solve multi-objective reliability optimization problems in this chapter. In this process, at each generation the following computation takes place: first the crowding distance is computed for selecting the global best particle and also for deleting particles from the external archive of nondominated solutions. The global best guide of the particles is selected from nondominated solutions with the highest crowding distance values. Selecting different guides for each particle in a specified top part of sorted repository based on a decreasing crowding distance allows the particles in the primary population to move toward those nondominated solutions in the external repository which are in the least crowded area in the objective space. Also, whenever archive is full, crowding distance is again used in selecting which solution to replace from the archive. This promotes diversity among the stored solutions in the archive since those solutions which are the most crowded areas are most likely to be replaced by a new solution. A flowchart showing the procedure of MOPSO-CD is provided in Fig. 1 [20,21].

Comparing with other evolutionary approaches, it has the following advantages [21–24]:

- (i) less parameters
- (ii) easy implementation
- (iii) fast convergence

3 Application of MOPSO-CD in reliability optimization problems

To evaluate the performance of MOPSO-CD for reliability optimization problems, a reliability test problem, namely series system is solved.

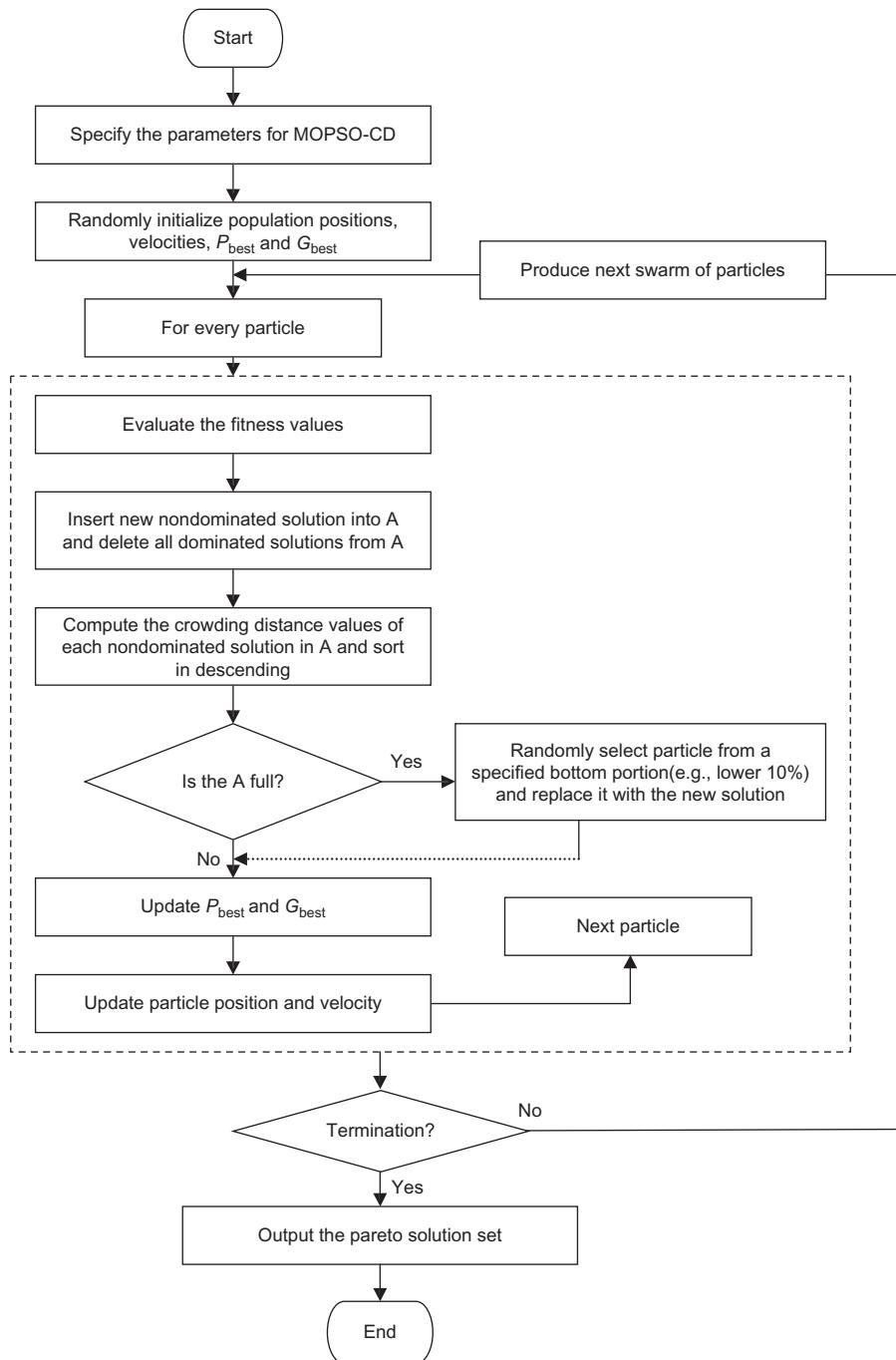


Fig. 1 The flow chart of MOPSO-CD.



Fig. 2 Block diagram of series system.

3.1 Series system

Suppose a certain series system as shown in Fig. 2 consists of five components, each having component reliability $r_j, j = 1, 2, 3, 4, 5$. The cost of j th component is assumed to be an increasing function of r_j in the form

$$c_j = a_j \log\left(\frac{1}{1-r_j}\right) + b_j, j = 1, 2, \dots, 5 \quad (3)$$

where vectors of coefficients a_j and b_j are $a = (24, 8, 8.75, 7.14, 3.33)$ and $b = (120, 80, 70, 50, 30)$, respectively. Thus the system reliability R_S , system unreliability Q_S , and system cost C_S are given by

$$R_S = \prod_{j=1}^5 r_j \quad (4)$$

or

$$Q_S = (1 - R_S) = 1 - \prod_{j=1}^5 r_j \quad (5)$$

$$C_S = \sum_{j=1}^5 c_j = a_j \log\left(\frac{1}{1-r_j}\right) + b_j \quad (6)$$

Then the MOOP problem is to determine the reliability of components, which maximizes the system reliability and minimizes the system cost or in other words, which minimizes both system unreliability and system cost. Thus the mathematical formulation of the problem is:

$$\begin{aligned} & \text{Find } (r_1, r_2, r_3, r_4, r_5) \text{ to minimize } (Q_S, C_S) \\ & \text{subject to} \\ & 0 \leq r_j \leq 1, \quad j = 1, 2, 3, 4, 5 \end{aligned} \quad (7)$$

Huang [25] solved this problem using fuzzy approach and reported only three Pareto-optimal solutions using aggregation method. We applied MOPSO-CD to solve this problem. The Pareto optimal front obtained by MOPSO-CD is shown in Fig. 3. The parameters used to solve the problem are shown in Table 1. The different solutions represent different combinations of system reliability and cost. The user has the flexibility and choice to choose a particular combination depending on the needs and constraints with the aid of a proper decision making process.

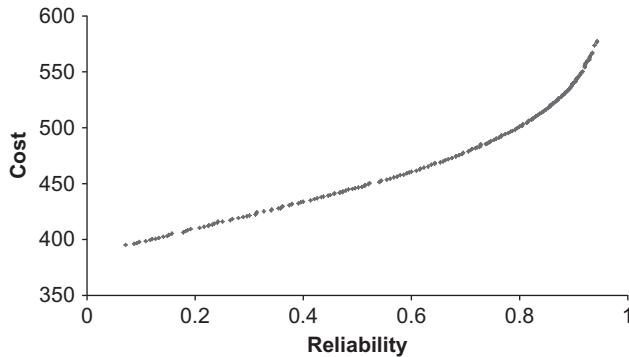


Fig. 3 Pareto front obtained by MOPSO-CD for series system.

Table 1 Parameters used in MOPSO-CD for series system

Pop size	Max gen	Mutation prob.	Archive size	c_1	c_2	w
200	300	0.25	220	1	1	0.3

4 Analysis of MOPSO-CD in reliability optimization problems

In this section, authors have analyzed the results obtained from MOPSO-CD for different parameter settings in reliability optimization problems of MOOP category. We have tried to check the different parameters on the optimization process. For each problem, first inertia weight is kept constant, and then the quality of Pareto optimal fronts for different mutation probabilities and acceleration coefficients is checked.

4.1 Series system

We have checked different settings of parameters and reported one of the best Pareto optimal fronts found during the investigation in Fig. 3. For further investigation we have tried to check the impact of different values of inertia weight, acceleration coefficient, and probability of mutation on the behavior of Pareto optimal front. Following three types of investigations have been made:

1. The Pareto optimal fronts for $w = 0.3, 0.6, 0.9$, and 1.20 have been checked.
2. The acceleration coefficients are kept at 1 (as suggested in original MOPSO-CD) and 2.0 (as suggested in most of the previously reported literature of PSO) for each of the aforementioned inertia weight.
3. For each above values of inertia, weight, and different acceleration coefficients, the probability of mutation has been tested for 0.3, 0.6, and 0.9.

For the series system first we have kept $w = 0.3$ and acceleration coefficients c_1 and c_2 are kept at 1.0. Then we kept changing the probability of mutation. For the different probabilities of mutation we have got almost same Pareto optimal fronts. It is further observed that these Pareto optimal fronts have good diversity and different probabilities of mutation provided almost same Pareto optimal fronts. These Pareto optimal fronts are shown in Figs. 4–6. But when we change acceleration coefficients c_1 and c_2 to the value 2.0 for the same value of inertia weight and probability of mutation, the Pareto optimal fronts lose their diversity and uniformity. These Pareto optimal fronts are shown in Figs. 7–9. Same experiments have been carried out for other

Fig. 4 Pareto optimal front for series system for $w = 0.3$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.3$.

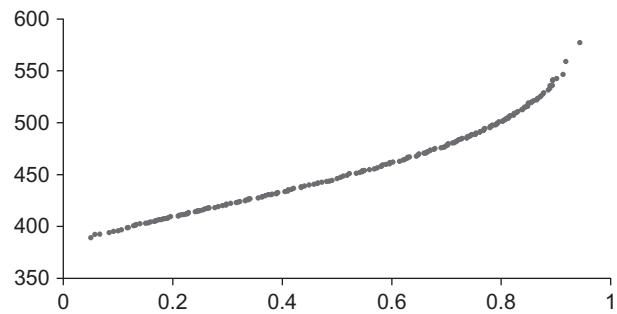


Fig. 5 Pareto optimal front for series system for $w = 0.3$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.6$.

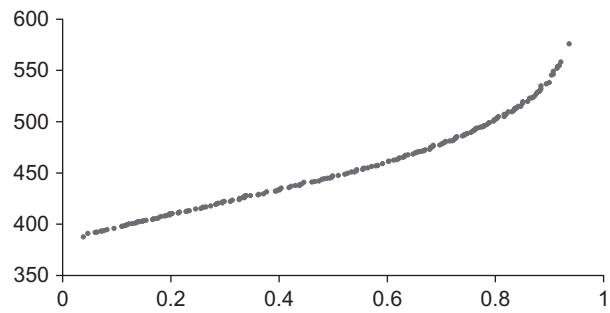
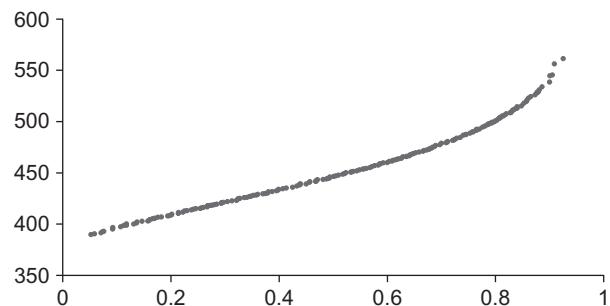


Fig. 6 Pareto optimal front for series system for $w = 0.3$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.9$.



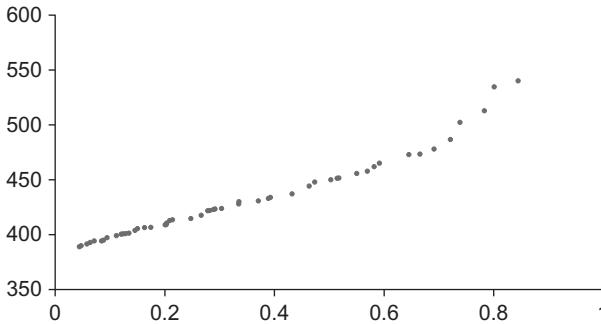


Fig. 7 Pareto optimal front for series system for $w = 0.3, c_1 = c_2 = 2.0$, and $P_{mut} = 0.3$.

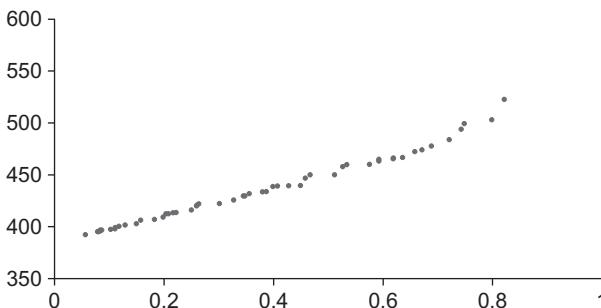


Fig. 8 Pareto optimal front for series system for $w = 0.3, c_1 = c_2 = 2.0$, and $P_{mut} = 0.6$.

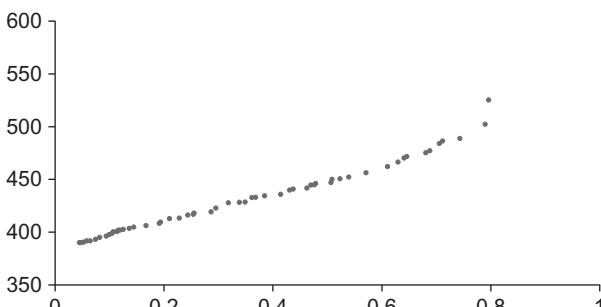


Fig. 9 Pareto optimal front for series system for $w = 0.3, c_1 = c_2 = 2.0$, and $P_{mut} = 0.9$.

different settings of parameters. Similarly many Pareto optimal fronts can be obtained varying the different parameters. When w is fixed at 0.6 and kept changing the probability of mutation and acceleration coefficients, we have got the same results. Although the Pareto optimal fronts for $w = 0.6$ are not as much good as for $w = 0.3$. Similar results are found for $w = 0.9$ and $w = 1.20$. These Pareto optimal fronts are shown in Figs. 10–27. It is clear from these plots that as we increase inertia weight from $w = 0.3$ to $w = 1.2$, diversity also decreases and at the same time by

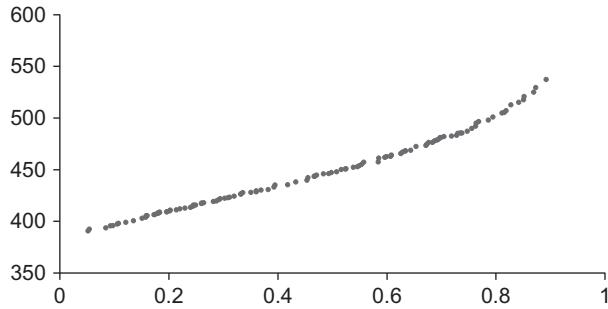


Fig. 10 Pareto optimal front for series system for $w=0.6, c_1=c_2=1.0$, and $P_{mut}=0.3$.

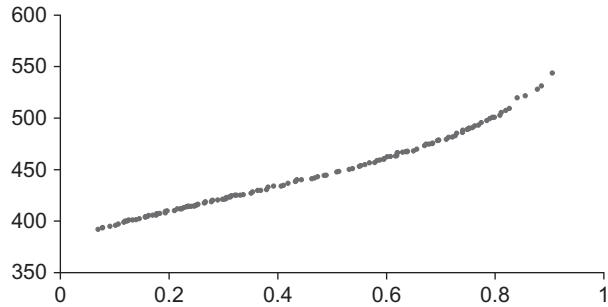


Fig. 11 Pareto optimal front for series system for $w=0.6, c_1=c_2=1.0$, and $P_{mut}=0.6$.

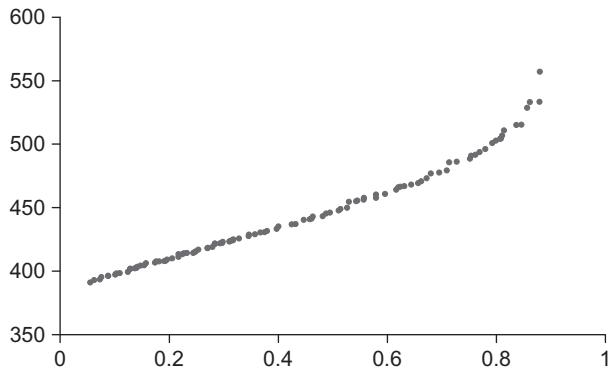


Fig. 12 Pareto optimal front for series system for $w=0.6, c_1=c_2=1.0$, and $P_{mut}=0.9$.

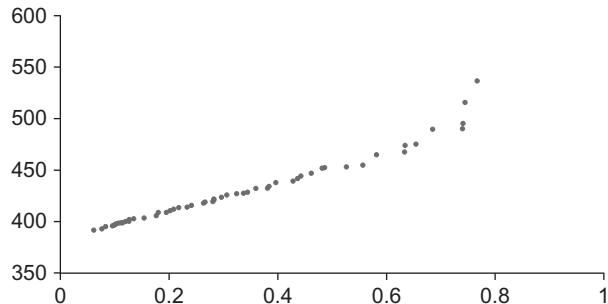


Fig. 13 Pareto optimal front for series system for $w=0.6, c_1=c_2=2.0$, and $P_{mut}=0.3$.

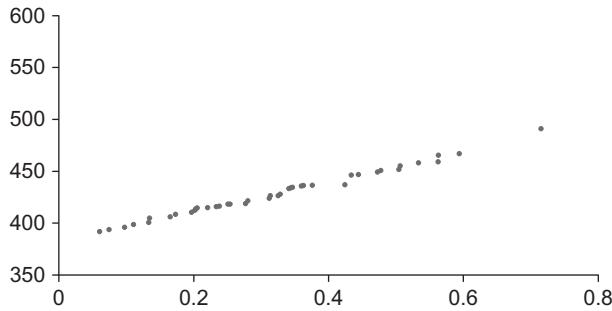


Fig. 14 Pareto optimal front for series system for $w = 0.6$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.6$.

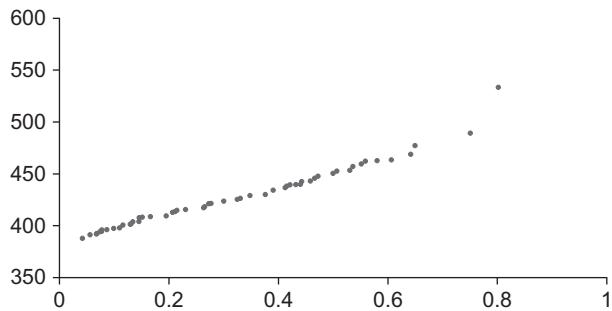


Fig. 15 Pareto optimal front for series system for $w = 0.6$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.9$.

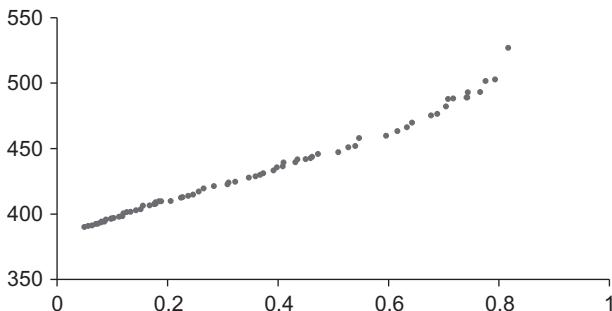


Fig. 16 Pareto optimal front for series system for $w = 0.9$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.3$.

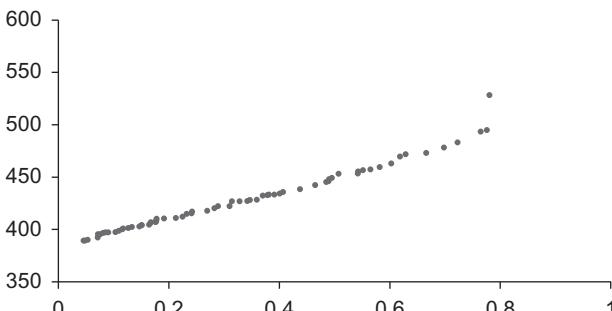


Fig. 17 Pareto optimal front for series system for $w = 0.9$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.6$.

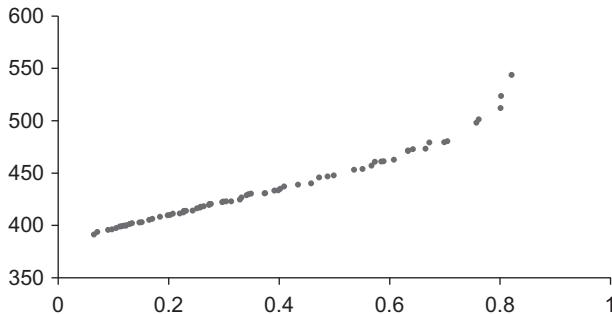


Fig. 18 Pareto optimal front for series system for $w=0.9$, $c_1=c_2=1.0$, and $P_{mut}=0.9$.

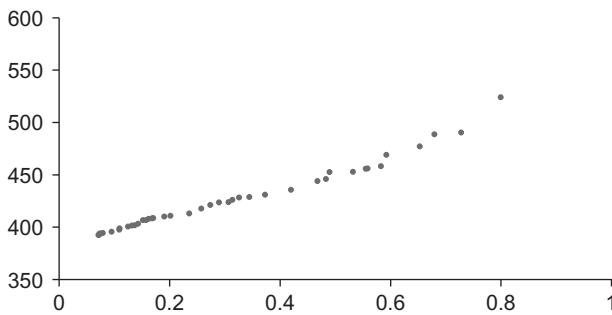


Fig. 19 Pareto optimal front for series system for $w=0.9$, $c_1=c_2=2.0$, and $P_{mut}=0.3$.

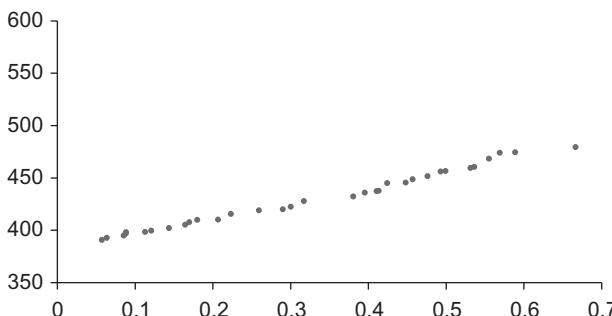


Fig. 20 Pareto optimal front for series system for $w=0.9$, $c_1=c_2=2.0$, and $P_{mut}=0.6$.

increasing value of acceleration coefficients for the different values of mutation probability the Pareto optimal fronts get poorer and lose their uniformity. Hence, for the series system the value of acceleration coefficient and inertia weight should not be very high as it can disturb the convergence and diversity of the algorithm.

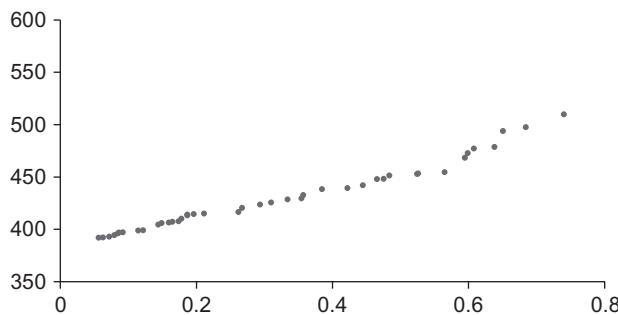


Fig. 21 Pareto optimal front for series system for $w = 0.9$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.9$.

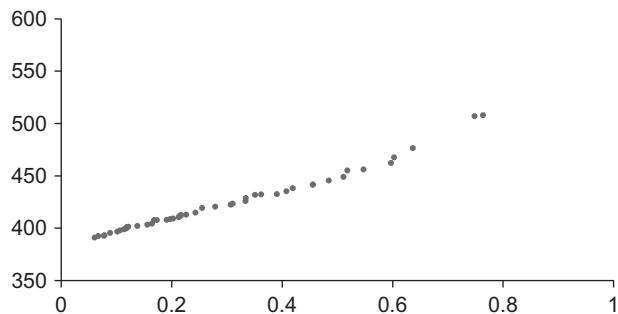


Fig. 22 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.3$.

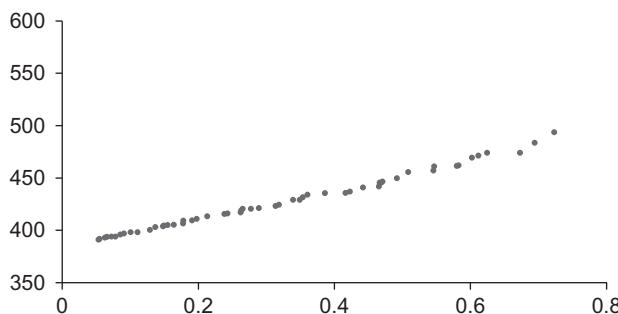


Fig. 23 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.6$.

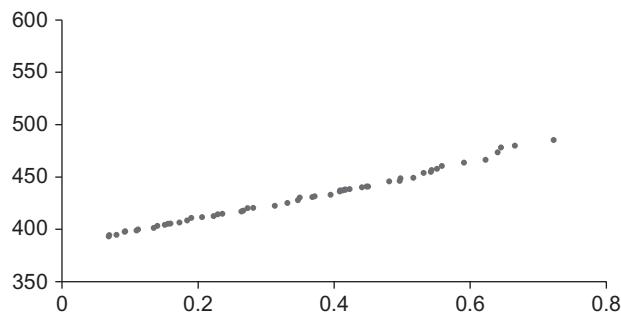


Fig. 24 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 1.0$, and $P_{mut} = 0.9$.

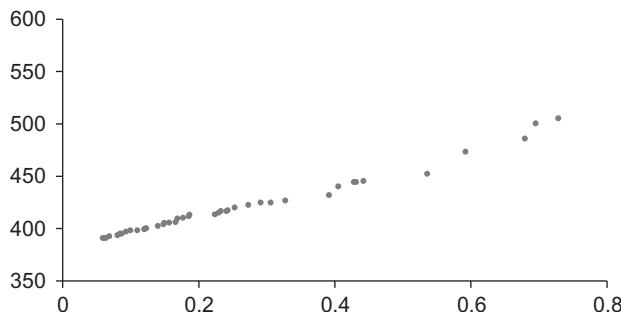


Fig. 25 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.3$.

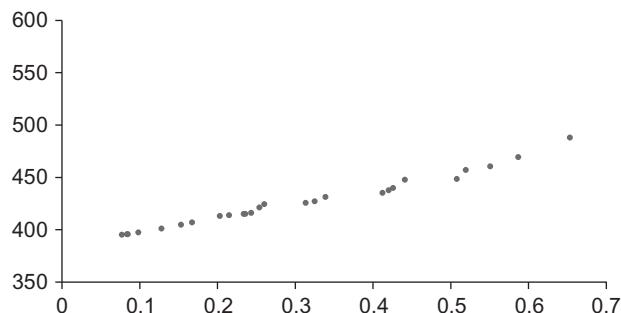


Fig. 26 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.6$.

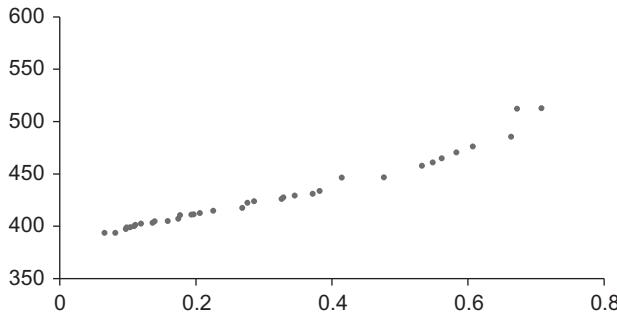


Fig. 27 Pareto optimal front for series system for $w = 1.2$, $c_1 = c_2 = 2.0$, and $P_{mut} = 0.9$.

5 Conclusion

Authors have applied MOPSO-CD for multi-objective reliability optimization problem. For this problem the MOPSO-CD algorithm evolves a good distribution of new Pareto-optimal solution set. The results are encouraging and suggest the applicability of the MOPSO-CD approach to more complex and real-world engineering problems. Apart from this, the MOPSO-CD is tested for different parameter settings. It is concluded that MOPSO-CD performed better for the lower value of acceleration coefficients and inertia weight but for higher values it can disturb the convergence and diversity of the algorithm.

Further, potential application of the well-proven technique MOPSO-CD in reliability optimization is presented. This theme links ideas and contributions that span the following areas:

- One of the most important practical advantages of the proposed PSO approach is that the mathematical models of real-life optimization problems can be solved.
- MOPSO-CD is examined for different settings of parameter, and it is concluded that inertia weight and acceleration coefficients have more effect on Pareto optimal front than the probability of mutation. The higher value (usually more than 0.6) of inertia weight is more responsible in lacking the diversity of swarm.
- The extension of the methodology especially to complex systems, and the results achieved, has demonstrated the efficiency of MPSO technique introduced in this work. The results obtained for reliability optimization problems are either better or comparable than those of previously reported results.

6 Future scope

- As part of future work plans, authors aim to develop a more efficient, optimized versions of the proposed algorithms so that it can tackle more complex problems. We have already identified some very interesting real-world problems coming from the wireless communication field (we have tried to optimize transmission/reception power, Antenna gain, date error rate,

- size, cost, etc. of devices and systems), which we could not tackle using the existing implementation due to too long run time.
- Also, MOPSO-CD is implemented only for reliability allocation problems, which are continuous nonlinear programming problem. The further work can be carried out to implement MOPSO-CD in redundancy allocation and reliability-redundancy allocation problems, which are integer nonlinear programming and mixed integer nonlinear programming problems, respectively.

References

- [1] A. Kumar, S.B. Singh, Reliability analysis of an n -unit parallel standby system under imperfect switching using copula, *Comput. Model. New Technol.* 12 (1) (2008) 47–55.
- [2] S. Pant, A. Kumar, A. Kishor, D. Anand, S.B. Singh, Application of a multi-objective particle swarm optimization technique to solve reliability optimization problem, in: Proceedings of IEEE Int. Conf. on Next generation Computing Technologies, September 4–5, 2015, pp. 1004–1007.
- [3] A. Kishore, S.P. Yadav, S. Kumar, Application of a multi-objective genetic algorithm to solve reliability optimization problem, in: Proceedings of IEEE International Conference on Computational Intelligence and Multi-Media Applications, 2007, pp. 458–462.
- [4] A.P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence*, John Wiley & Sons, Hoboken, New Jersey, USA, 2005.
- [5] B. Clow, T. White, An evolutionary race: a comparison of genetic algorithms and particle swarm optimization used for training neural networks, in: Proceedings of International Conference on Artificial Intelligence, 2004, pp. 582–588.
- [6] C.R. Raquel, P.C. Naval Jr., An effective use of crowding distance in multiobjective particle swarm optimization, in: Genetic and Evolutionary Computation Conference, 2005, pp. 257–264.
- [7] C.M. Rocco, A.J. Miller, J.A. Moreno, N. Carrasquero, A cellular evolutionary approach applied to reliability optimization of complex systems, in: Proceedings of the Annual Reliability and Maintainability Symposium, 2000, pp. 210–215.
- [8] H. Garg, M. Rani, S.P. Sharma, Y. Vishwakarma, Intuitionistic fuzzy optimization technique for solving multi-objective reliability optimization problems in interval environment, *Expert Syst. Appl.* 41 (7) (2014) 3157–3167.
- [9] H. Garg, S.P. Sharma, Multi-objective reliability-redundancy allocation problem using particle swarm optimization, *Comput. Ind. Eng.* 64 (1) (2013) 247–255.
- [10] S. Guangyong, L. Guanyao, G. Zhihui, C. Xiangyang, Y. Xujing, L. Qing, Multiobjective robust optimization method for drawbead design in sheet metal forming, *Mater. Des.* 31 (4) (2010) 1917–1929.
- [11] S. Pant, S.B. Singh, Particle swarm optimization to reliability optimization in complex system, in: IEEE Int. Conf. on Quality and Reliability, Bangkok, Thailand, September 14–17, 2011, pp. 211–215.
- [12] S. Pant, D. Anand, A. Kishor, S.B. Singh, A particle swarm algorithm for optimization of complex system reliability, *Int. J. Perform. Eng.* 11 (1) (2015) 33–42.
- [13] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.
- [14] A. Kumar, S. Pant, S.B. Singh, Reliability optimization of complex system by using cuckoos search algorithm, *Mathematical Concepts and Applications in Mechanical*

- Engineering and Mechatronics, IGI Global Publisher, Hershey, Pennsylvania, USA, 2017, pp. 94–110.
- [15] M. Mutingi, V.P. Kommula, Reliability optimization for the complex bridge system: fuzzy multi-criteria genetic algorithm, in: Proceedings of Fifth International Conference on Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing, vol. 437, 2016, pp. 651–663, http://dx.doi.org/10.1007/978-981-10-0451-3_58.
 - [16] D. Salazar, C.M. Rocco, B.J. Galvan, Optimization of constrained multiple-objective reliability problems using evolutionary algorithms, *Reliab. Eng. Syst. Saf.* 91 (2006) 1057–1070.
 - [17] V. Ravi, B.S.N. Murty, J. Reddy, Nonequilibrium simulated-annealing algorithm applied to reliability optimization of complex systems, *IEEE Trans. Reliab.* 46 (2) (1997) 233–239.
 - [18] C.A.C. Coello, G.T. Pulido, M.S. Lechuga, Handling multiple objectives with particle swarm optimization, *IEEE Trans. Evol. Comput.* 8 (2004) 256–279.
 - [19] J. Safari, Multi-objective reliability optimization of series-parallel systems with a choice of redundancy strategies, *Reliab. Eng. Syst. Safe.* 108 (2012) 10–20.
 - [20] J. Zhao, Z. Liu, M. Dao, Reliability optimization using multi objective ant colony system approaches, *Reliab. Eng. Syst. Saf.* 92 (1) (2007) 109–120.
 - [21] R. Hassan, B. Cohanim, O. De Weck, G. Venter, A comparison of particle swarm optimization and the genetic algorithm, in: Proceedings of the 1st AIAA Multidisciplinary Design Optimization Specialist Conference, 2005, pp. 1–13.
 - [22] M.S. Chern, On the computational complexity of reliability redundancy allocation in a series system, *Oper. Res. Lett.* 11 (5) (1992) 309–315.
 - [23] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995, pp. 39–43.
 - [24] R.J.W. Hodgson, Particle swarm optimization applied to the atomic cluster optimization problem, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2002, pp. 68–73.
 - [25] H.Z. Huang, Fuzzy multi-objective optimization decision-making of reliability of series system, *Microelectron. Reliab.* 37 (1997) 447–449.

Free vibration and connecting bolt constraint-based FEA analysis of heavy vehicle medium duty transmission gearbox housing made from AISI 4130 alloy material

7

A. Kumar, P.P. Patil
Graphic Era University, Dehradun, India

1 Introduction

This research work signifies the material-based effect on natural frequencies and mode shapes. It is observed from the previous studies that internal excitation is the reason for transmission housing failure. In open literature survey various studies were examined to form the base of this research work. Dogan [1] has investigated to reduce noise and vibration effect in transmission gearbox. The undesirable noise of rattle and clattering is produced due to torsional vibration. For experimental analysis the transmission parameters were varied to reduce the effect of rattle and clatter noise. Wang and yang [2] have investigated the nonlinearity on gear tooth face. In this research work chaos, bifurcation with sliding friction was studied. Abouel-Seoud and Abdallah [3] have studied car gearbox using vibration response method. Vandi and Ravaglioli [4] in this paper present implementation of a simplified engine-driveline model to complete an existing vehicle dynamic model. The engaged and disengaged phenomena of clutch were investigated.

Nacib and Sakhara [5] have studied the heavy gearbox of helicopters. Gordon and Bareket [6] have studied the source of vibration in sports utility vehicle; they have used sensor and data acquisition system to find the sources of vibration. In sources of vibration they have also considered road surface features. Gear vibration, vibroacoustic analysis of automotive structures, and high-power toothed gears were performed using MCSA (motor current signature analysis) and DWT (discrete wavelet transform) method [7–10]. Yu and Xia [11] have used the finite element method for the structural optimization of the grey cast iron HT200 gearbox housing. Kuo [12] has investigated the engagement and disengagement of the joint elements for an automatic transmission powertrain using Matlab/Simulink.

Yulong and Weipeng [13] focused on a dual-clutch automatic transmission of its hydraulic system. Sayer and Busmann et al. [14] in this paper highlight the influence of Materials' mechanical properties. The present work investigates about change of

material properties by the influence of loads and environmental conditions. The authors' research of dynamic characteristics of transmission gearbox housing has significant impact on the reduction of vibration and noise. Grey cast iron has good damping properties to reduce the magnitude of vibration waves.

2 CAD model and material properties

3D solid model of heavy vehicle truck transmission gearbox housing consists of more than 600 parts. It consists of shafts, gears, mountings, etc. Solid Edge and Pro-E [15,16] were used as modeling tools suited for complex geometry. Free vibration analysis was performed and first twenty inherent natural frequencies and mode shapes were evaluated. The 3D CAD model of transmission gearbox housing is shown in Fig. 1. For free vibration analysis finite element-based software, Ansys 14.5 [17] was used. Fig. 2 shows the discretized FEA model of transmission gearbox housing.

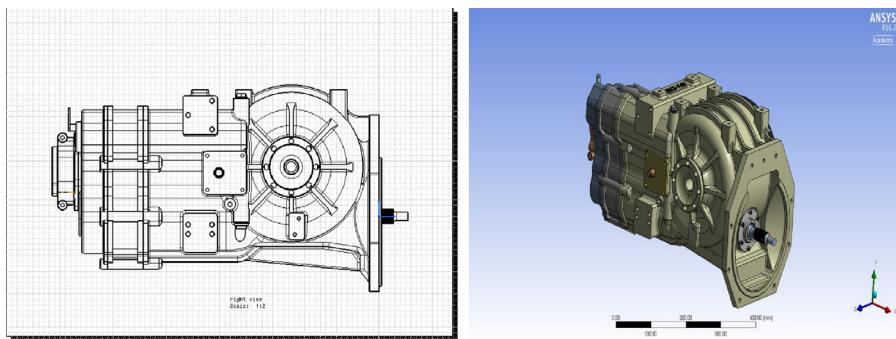


Fig. 1 3D solid model of transmission gearbox assembly.

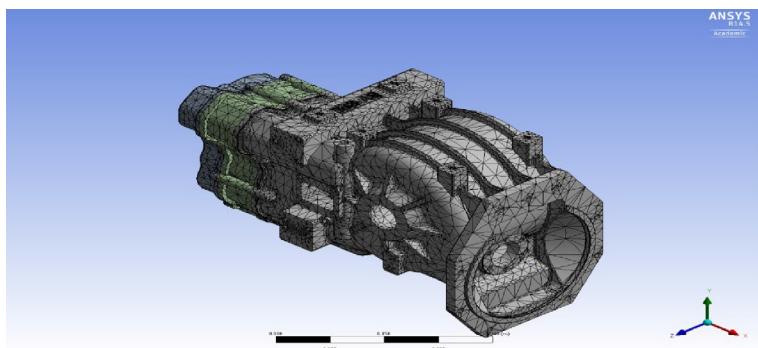


Fig. 2 FEA meshing of transmission housing.

The meshed model of AISI 4130 consists of 196,137 nodes and 113,566 elements. FEA meshing was performed using Linear Tetrahedron (Tet 4) elements.

Materials' mechanical properties, density, Young's modulus, and Poisson's ratio were required for numerical simulation of free vibration. Material properties of grey cast iron HT200 [11], grey cast iron FG260 [18], and AISI 4130 were used for FEA analysis. The mechanical properties of Grey Cast Iron HT200 are density 7200 kg/m³, Young's modulus 1.1 e11 Pa, Poisson's ratio 0.28. The mechanical properties of grey cast iron FG260 are density 7200 kg/m³, Young's modulus 1.28 e11 Pa, and Poisson's ratio 0.26. The mechanical properties of AISI 4130 are density 7861.1 kg/m³, Young's modulus 2.05 e11 Pa, Poisson's ratio 0.26, bulk modulus 1.4236 e11 Pa, and shear modulus 8.1349 e10 Pa. Density of AISI 4130 and structural steel material is in same range [19].

3 FEA simulation: modal analysis

Free vibration analysis in FEA deals with coordinate's transformation. Natural coordinates replace nodal coordinates and for decoupling motion equations physics coordinates replace modal coordinates. In linear analysis, second order differential equation is solved as follows:

$$[M]\{\ddot{U}\} + [C]\{\dot{U}\} + [K]\{U\} = \{F(t)\} \quad (1)$$

where

$[M]$, $[C]$, and $[K]$ are mass, damping, and stiffness matrices.

$\{U\}$, $\{\dot{U}\}$, and $\{\ddot{U}\}$ are displacement, speed, and acceleration time-dependent vectors.

$\{F(t)\}$ is the external force vector.

The external force consists of harmonic excitation that is independent of structural response therefore neglecting the external force for modal analysis and considering undamped structure.

$$F(t) = 0, \quad C = 0$$

$$[M]\{\ddot{U}\} + [K]\{U\} = 0 \quad (2)$$

Eq. (2), homogenous differential equation, is solved using harmonic solution:

$\{U(t)\} = \{\Psi\} \exp(j\Omega t)$, ($j^2 = -1$) to find eigenvalues and eigenvectors.

$[\Omega^2]$, and $[\Psi]$ are matrices of natural frequencies and corresponding mode shapes.

4 FEA results and discussion

FEA-based modal analysis of heavy vehicle medium duty transmission gearbox housing evaluates result for first twenty inherent natural frequencies and mode shape. Zero displacement constraint-based boundary condition was used for the simulation. FEA

simulation shows that the natural frequencies of grey cast iron HT200 varies (1669–3576 Hz), grey cast iron FG260 varies (1801.4–3852.6 Hz), and AISI 4130 alloy material (2157–4659.9 Hz). The higher order frequency variation shows the excellent structural rigidity by eliminating lower order frequency. The critical frequency range for heavy vehicle truck transmission housing varies (0–74 Hz). So in order to eliminate the resonance, fundamental frequencies should not fall in this range.

In numerical simulation of grey cast iron HT200 various vibration modes such as torsional vibration and axial bending vibration were identified. Torsional vibration and axial bending cause more harms to transmission housing. Mode 1 shows heavy vibration and deformation at the center of housing. Torsional vibration was performed at the center side of transmission housing. Mode 19 shows heavy vibration at center and left corner with deformations. Axial bending vibration was found in modes 9, 11, and 15. Mode 20 has the highest natural frequency of 3576 Hz with lowest deformation. The overall weight of grey cast iron HT200 transmission gearbox housing is 54.85 kg.

In numerical simulation for grey cast iron FG260, modes 1, 5, and 8 are torsional. Modes 1 and 5 show deformation at center side. Mode 8 shows deformation in bottom portion. Axial bending vibrations were found in modes 9, 11, and 12. Modes 15, 17, and 19 are axial bending vibration with torsional vibration; both axial bending and torsional vibration happen in upper and lower sides. The overall weight of grey cast iron HT200 transmission gearbox housing is 54.85 kg.

Selected mode shapes and corresponding frequencies are shown in Fig. 3. Modes 1, 3, and 5 show normal level of deformations with less vibration effects. Mode 6 has torsional vibration having maximum deformation at center portion (red hues). Mode 13 shows vibration effect at the extreme left corner of housing having large deformations. Mode 20 shows axial bending vibrations about x-axis but has very less effect on housing due to high structural rigidity. The calculated mass of gearbox housing using AISI 4130 material is 63.711 kg that is 8.861 kg higher in weight.

5 FEA natural frequency comparison for different materials

Fig. 4 shows graphical comparison of grey cast iron and AISI 4130 frequency variations. Grey cast iron HT200 and FG260 both have approximate same material properties and are well suited for transmission housing material. Due to the damping effect and casting properties, these two materials are preferred for manufacturing of transmission housing. AISI 4130 is recommended for housing material based on the present study.

In the present study a new material AISI 4130 has been suggested for future application. AISI 4130 suitability for transmission housing is analyzed on design and vibration index. The manufacturing and cost-effectiveness index was not studied in this paper. AISI 4130 alloy steel has chromium and molybdenum as strengthening agents. It is more responsive to heat treatments. Grey cast iron housing is manufactured using casting process in two halves. Due to low carbon content AISI 4130 has excellent weldability.

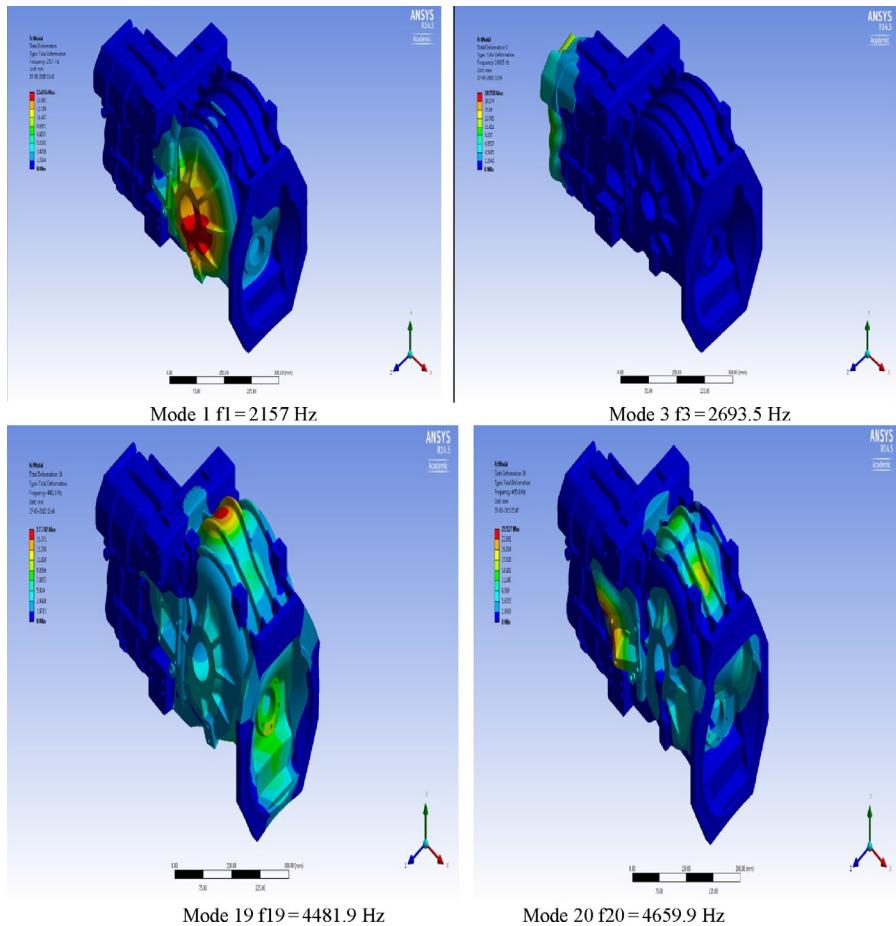


Fig. 3 Mode shape and natural frequency of AISI 4130 transmission gearbox housing.

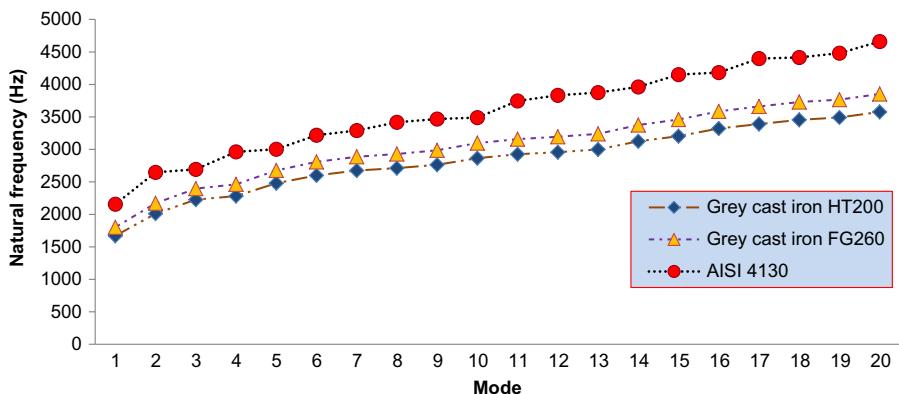


Fig. 4 Frequency variations for AISI 4130 and grey cast iron materials.

To study vibration index zero displacement constraint-based boundary condition was applied. The modal frequency of AISI 4130 transmission housing varies (2157–4659.9 Hz). This range of frequency is higher in comparison to grey cast iron HT200 and FG260. It is much higher than critical frequency range (0–74 Hz). AISI 4130 material range of frequency matches with safe frequency of Tuma [10] results. Tuma experimental frequency range varies (500–3000 Hz) and heavy vibration occurs in (500–2500 Hz). The chances of heavy vibration and resonance are eliminated due to higher frequency (2157–4659.9 Hz), which shows excellent structural rigidity with higher strength. On vibration index AISI 4130 material suited for heavy vehicle medium duty truck transmission housing.

In Fig. 4 graphical comparison of FEA simulation results has been done between grey cast iron and AISI 4130. Grey cast iron HT200 and FG260 both have approximate same material properties and are well suited for transmission housing material. Due to damping effect and casting properties these two materials are preferred for manufacturing of transmission housing. AISI 4130 is recommended for housing material based on the present study. Fig. 4 illustrates that all material frequencies are in higher range. This condition prevents resonance and heavy vibration of transmission system.

6 Connecting bolt constraint analysis

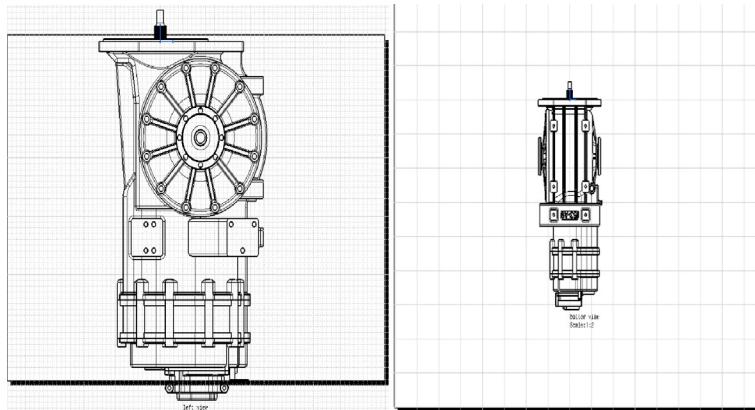
In the first stage of FEA analysis it was concluded from FEA results that AISI 4130 material is suitable for transmission housing. In the second stage connecting bolt (CB) constraint analysis was performed on AISI 4130 transmission housing. In general transmission housing is consolidated on vehicle frame using CBs. These CBs constrain the motion of housing in all directions and prevent looseness and harmonic excitation conditions. Housing is fixed on frame using all 37 CBs; this condition is known as zero displacement condition and FEA analysis is known as constraint-based analysis. Number of CBs used for constraining may vary depending upon the shape and size of the transmission housing.

When housing is consolidated on vehicle frame using 37 bolts in all five directions, natural frequency of housing varies (2157–4659.9 Hz). In this research work varying number of CBs (Table 1) has investigated the effect of CBs on fundamental and natural frequencies. Fig. 5 shows position of CBs in housing. Side position shows 7 and upper position shows 6 CBs constraint position. For actual running simulation in FEA these positions of CBs were constrained.

First twenty natural frequencies are shown in Table 1. For six CB constraint condition natural frequency vary (1053.4–2941.2 Hz). Fundamental frequency 1053.4 Hz is in lower range that may cause resonance condition. Resonance in heavy vehicle transmission housing leads heavy vibration and failure of transmission components. To prevent this condition CBs number has been increased. When transmission housing was consolidated on vehicle frame using 8 CBs, natural frequency varies (1447.6–3087.2 Hz). Fundamental frequency 1447.6 Hz is in higher order range which prevents heavy vibration. Using 12 CBs natural frequency shifted in higher

Table 1 Frequency variation for connecting bolts variation

S. No.	6 Connecting bolts	8 Connecting bolts	12 Connecting bolts
1	1053.4	1447.6	1547.9
2	1171.5	1463	1681.5
3	1240.1	1483.6	1776.7
4	1377.1	1569.2	1986
5	1538.8	1764.3	2152.8
6	1599.4	1825.9	2229
7	1780.6	1914.8	2335.4
8	1850.3	1978.8	2376.1
9	1907.2	2013.4	2503
10	1975.2	2132.1	2631.6
11	2080	2407.6	2714.5
12	2133.6	2425.4	2823.8
13	2205.9	2567.4	2832.7
14	2291.8	2610.7	3004.3
15	2404.7	2659.8	3027.1
16	2514.5	2697.9	3104
17	2583.4	2832.4	3111.8
18	2648.9	2889.5	3216.4
19	2744.6	3020.5	3260.1
20	2941.2	3087.2	3349.6

**Fig. 5** Connecting bolts positions in transmission housing.

range (1547.9–3349.6 Hz). From CBs optimization results the use of 12 number of CBs was suggested to consolidate housing on vehicle frame due to prevention from torsional vibration twisting action.

Fig. 6 shows graphical comparison of natural frequency. It is shown in graph that as the number of CBs increases, frequency increases linearly. In FEA simulation first

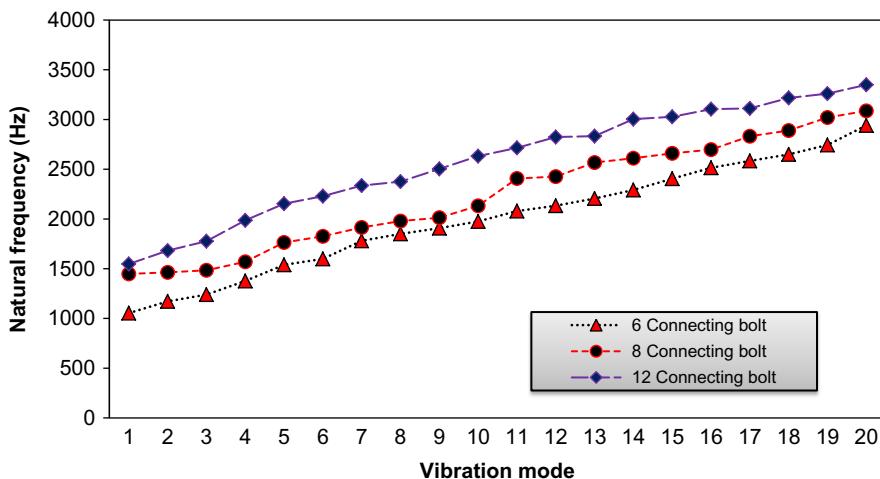


Fig. 6 Graphical comparison of frequency variation.

20 natural frequency and mode shapes were evaluated using 6, 8, and 12 CBs constraint conditions. [Figs. 7–9](#) show mode shape analysis for deformation and vibration effect.

[Fig. 7](#) shows vibration mode shapes of transmission housing when housing is fixed using 6 CBs. Vibration mode shapes of AISI 4130 heavy vehicle medium duty transmission housing were evaluated using modal analysis. Mode 1 shows small deformation region on front side of housing in red hues. Mode 5 shows effect of torsional vibration performing at the end side of transmission housing and shows (15–17) mm of total deformation. Modes 12 and 16 show axial bending with torsional vibration. This condition shows maximum deformation in transmission housing. Mode

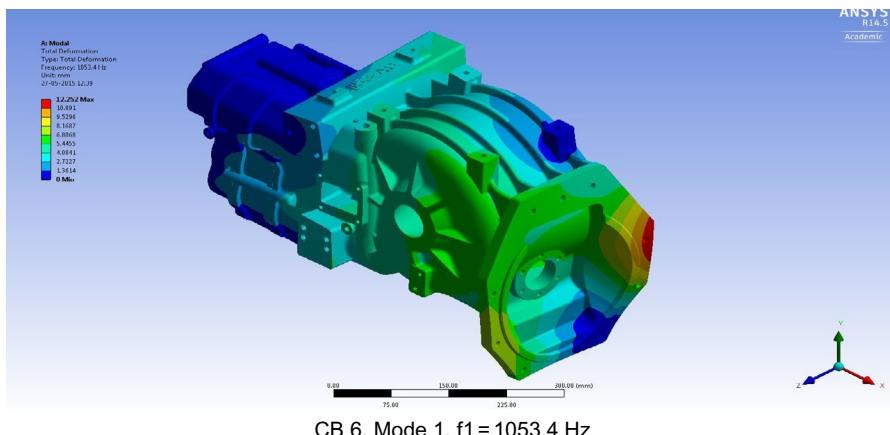
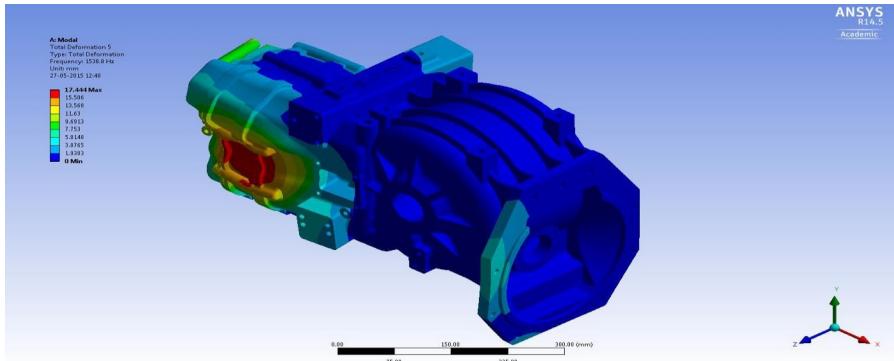
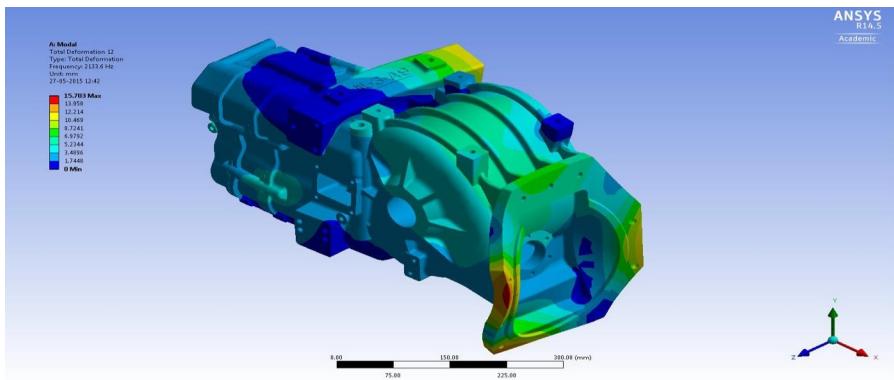
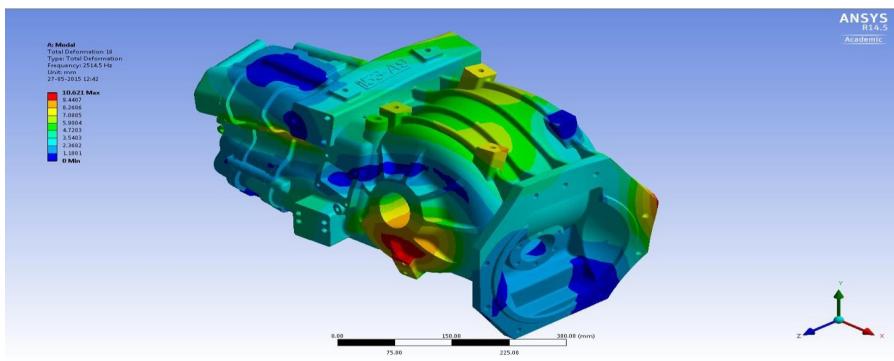


Fig. 7 Six connecting bolt-based different mode shapes and frequencies.

(Continued)

CB 6, Mode 5, $f_5 = 1538.8$ HzCB 6, Mode 12, $f_{12} = 2133.6$ HzCB 6, Mode 16, $f_{16} = 2514.5$ Hz**Fig. 7, cont'd**

16 shows higher deformation at center portion of transmission housing (red hues). Higher deformation and more effect of axial and torsional vibrations in the main region due to 6 CB constraint-based condition are not recommended for the consolidation of transmission housing on vehicle frame.

Fig. 8 shows vibration mode shapes of housing constraint by 8 CBs. Mode 2 shows axial bending vibration along the center line. Back portion of transmission housing is subjected to (7–8 mm) high deformation. Mode 8 shows vibration effect at down side of housing. Mode 12 shows effect of torsional vibration on transmission housing. Housing front end is twisted and deformed. Mode 16 shows torsional and axial bending vibration combined effect. Whole housing is subjected to deformation and heavy

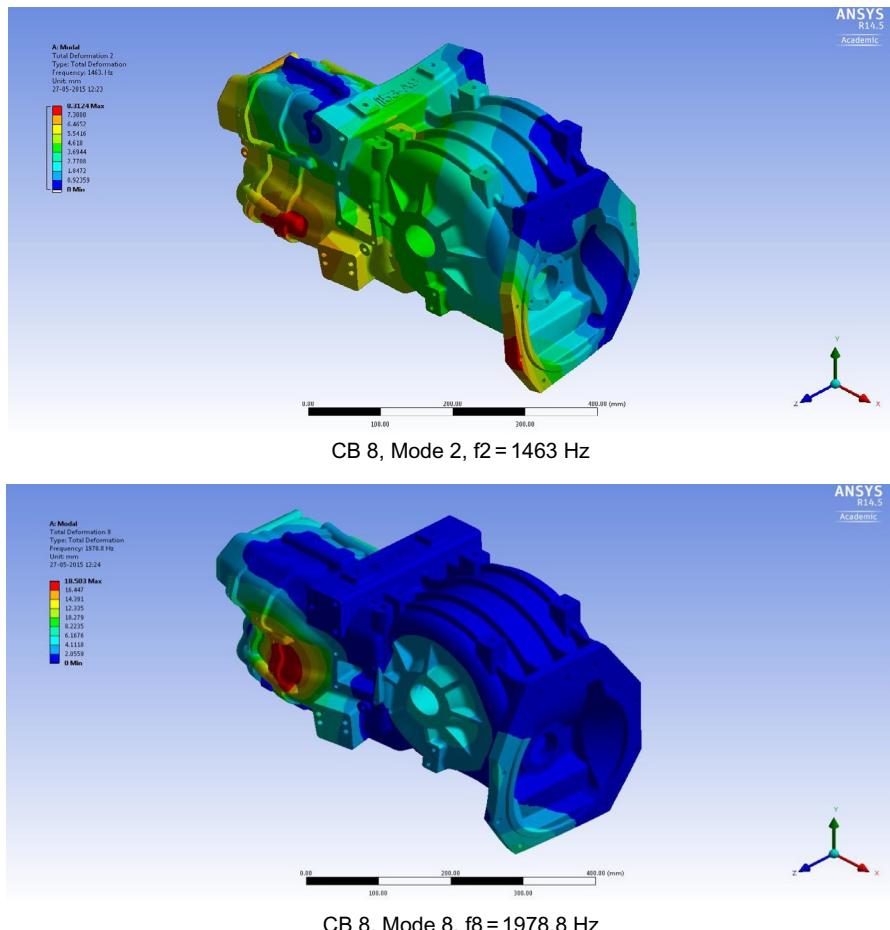


Fig. 8 Eight connecting bolts-based different mode shapes and frequencies.

(Continued)

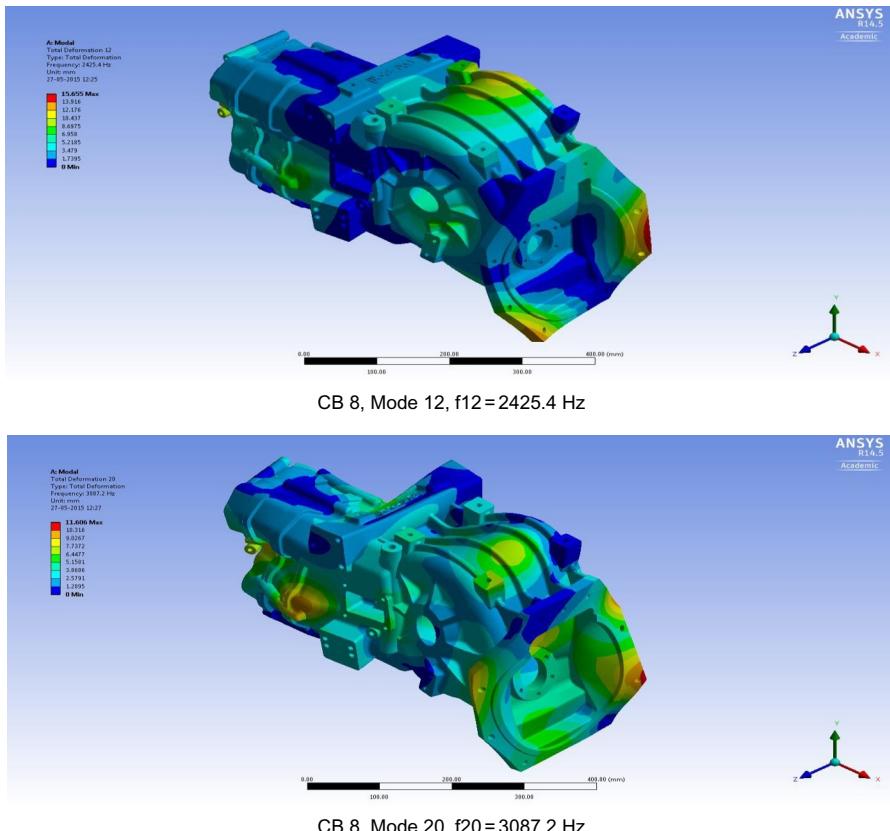
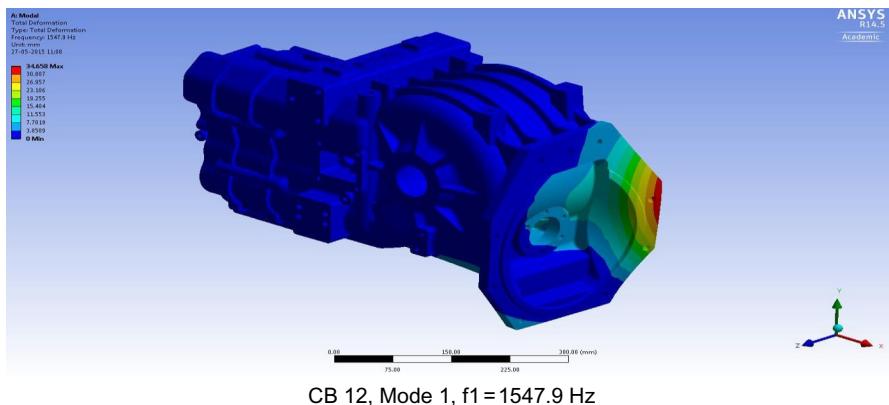


Fig. 8, cont'd

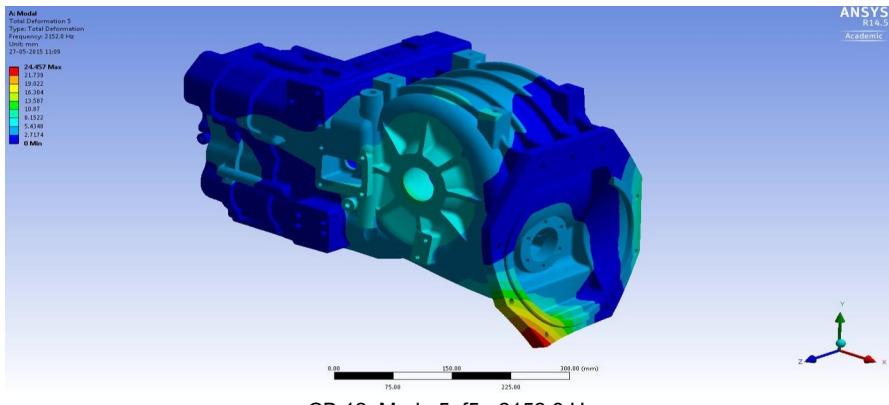
vibration effect. This is the prime reason for eliminating 8 CBs constraint condition even when fundamental frequency is in higher order range.

CBs 6 and 8-based constraint analysis shows higher effect of torsional and axial bending vibration. These two vibrations (torsional and axial bending vibration) cause heavy vibration with large deformation in housing. To find reduced vibration condition 12 CBs-based constraint condition was applied.

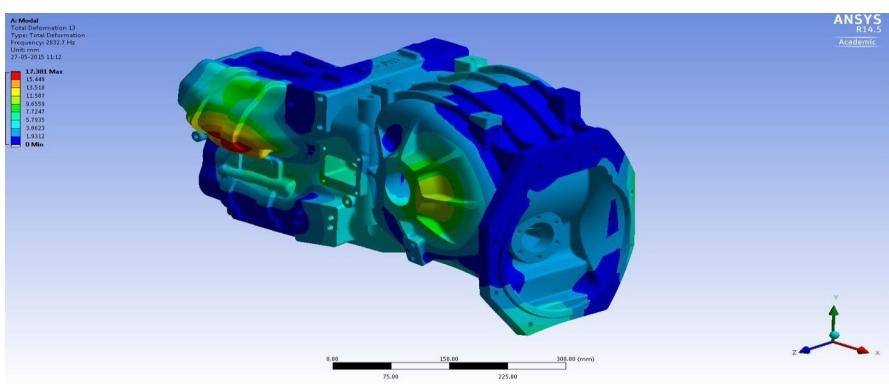
Fig. 9 shows vibration mode shapes of housing when it is consolidated on vehicle frame using 12 CBs. Mode 1 shows no higher deformation regions at fundamental frequency 1547.9 Hz. Mode 5 shows small deformed region at front corner end of transmission housing (red hues). Mode 13 shows torsional vibration at extreme corner of housing. Mode 15 shows normal vibration condition. It was concluded from the mode shapes that when housing was consolidated on vehicle frame using 12 CBs housing is in safe condition due to less deformation and less effect of torsional and axial bending vibrations.



CB 12, Mode 1, f1 = 1547.9 Hz



CB 12, Mode 5, f5=2152.8 Hz



CB 12, Mode 13, f13 = 2832.7 Hz

Fig. 9 Twelve connecting bolt-based different mode shapes and frequencies.

(Continued)

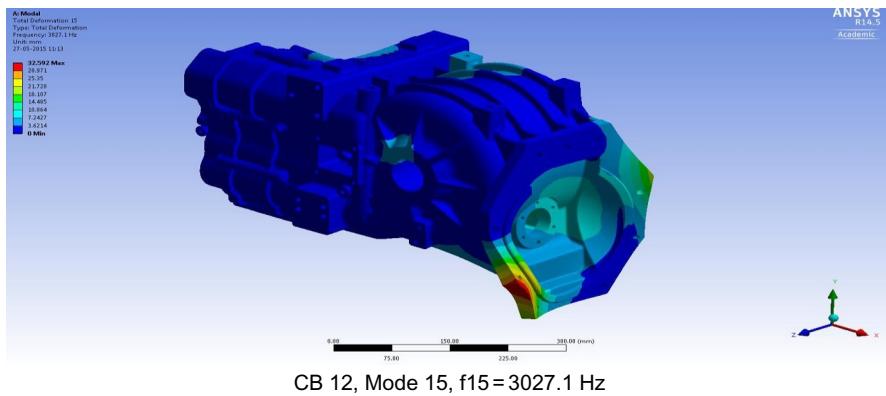


Fig. 9, cont'd

7 Conclusions

This research work has significance in design stage of heavy vehicle transmission gearbox housing. FEA-based numerical simulation was performed in two stages for thorough analysis of AISI 4130 transmission housing. The concluding remarks from the present study are as following:

- Weight of housing was calculated for three different materials (grey cast iron HT200, grey cast iron FG260, and AISI 4130). On design and vibration index, all three materials can be used as a truck transmission casing/housing.
- A new alloy steel AISI 4130 having chromium and molybdenum as chief elements for strength enhancement is recommended for future application of transmission gearbox housing based on design and vibration index.
- It was found that 12 CB constraint-based conditions reduce vibration and deformation in AISI 4130 transmission housing. So it was concluded that 12 CBs can be used for constraining housing on vehicle frame.

Acknowledgment

Authors are thankful to Department of Science and Technology (DST) New Delhi and research cell of Graphic Era University (GEU), Dehradun for necessary funding of advanced modeling and simulation lab where this research work was carried out.

References

- [1] S.N. Dogan, Loose part vibration in vehicle transmissions—gear rattle, *Trans. J. Eng. Environ. Sci.* 23 (1999) 439–454.
- [2] J. Wang, J. Zheng, A. Yang, An analytical study of bifurcation and chaos in a spur gear pair with sliding friction, *Procedia Eng.* 31 (2012) 563–570.

- [3] S.S. Abouel-Seoud, E.S. Mohamed, A.A. Abdel-Hamid, A.S. Abdallah, Analytical technique for predicting passenger car gearbox structure noise using vibration response analysis, *Br. J. Appl. Sci. Technol.* 3 (4) (2013) 860–883.
- [4] G. Vandi, N. Cavina, E. Corti, G. Mancini, D. Moro, F. Ponti, V. Ravaglioli, Development of a software in the loop environment for automotive powertrain systems, *Energy Procedia* 45 (2014) 789–798.
- [5] L. Nacib, K.M. Pekpe, S. Sakhara, Detecting gear tooth cracks using cepstral analysis in gearbox of helicopters, *Int. J. Adv. Eng. Technol.* 5 (2013) 139–145.
- [6] T.J. Gordon, Z. Bareket, Vibration transmission from road surface features—vehicle measurement and detection. Technical Report for Nissan Technical Center North America, Inc. UMTRI-2007-4, 2007.
- [7] C. Kar, A.R. Mohanty, Monitoring gear vibrations through motor current signature analysis and wavelet transform, *Mech. Syst. Signal Process.* 20 (2006) 158–187.
- [8] P. Czech, Diagnosis of industrial gearboxes condition by vibration and time frequency, scale-frequency, frequency-frequency analysis, *Metalurgija* 51 (2012) 521–524.
- [9] A. Kumar, P.P. Patil, FEA simulation based performance study of multi-speed transmission gearbox, *Int. J. Manufac. Mater. Mech. Eng.* 6 (1) (2016) 57–67, <http://dx.doi.org/10.4018/IJMMME.2016010103>.
- [10] J. Tuma, Gearbox noise and vibration prediction and control, *Int. J. Acoust. Vib.* 14 (2009) 1–11.
- [11] F. Yu, Y. Li, D. Sun, W. Shen, W. Xia, Analysis for the dynamic characteristic of the automobile transmission gearbox, *Res. J. Appl. Sci. Eng. Technol.* 5 (2013) 1449–1453.
- [12] K. Kuo, Simulation and analysis of the shift process for an automatic transmission, *World Acad. Sci. Eng. Technol.* 52 (2011) 341–347.
- [13] L. Yulong, L. Xingzhong, L. Weipeng, Hydraulic system optimization and dynamic characteristic simulation of double clutch transmission, *Procedia Environ. Sci.* 10 (2011) 1065–1070.
- [14] F. Sayer, F. Bürkner, M. Blunk, A.M. Wingerde, H.G. Busmann, Influence of loads and environmental conditions on material properties over the service life of rotor blades, *Dewi Magazin No. 34* (2009) 24–31.
- [15] Solidedge, Version 19.0, 2006.
- [16] Pro-E 5.0, Designing Guide Manual, 2013.
- [17] Ansys R 14.5, Academic, Structural Analysis Guide, 2013.
- [18] A. Kumar, A. Dwivedi, H. Jaiswal, P.P. Patil, Material based vibration characteristic analysis of heavy vehicle transmission gearbox casing using finite element analysis (FEA), *Adv. Int. Syst. Comput.* 308 (2015) 527–533.
- [19] A. Kumar, H. Jaiswal, P.P. Patil, FEM simulation based computation of natural frequencies and mode shapes of loose transmission gearbox casing, *Int. Rev. Model. Simulat.* 7 (2014) 900–905. <http://dx.doi.org/10.15866/iremos.v7i5.3932>.

Maintenance grouping models for multicomponent systems

8

P. Do^{*}¹, A. Barros[†]

*University of Lorraine, Vandoeuvre-les-Nancy, France, [†]Norwegian University of Science and Technology, Trondheim, Norway

¹Corresponding author: e-mail address: phuc.do@univ-lorraine.fr

1 Introduction

In the framework of reliability theory and stochastic modeling, the system to be maintained is modeled from a functional point of view, that is to say we look at the way a main function is fulfilled. Then the subsystems and components are described in light of this main function, to express how they interact and can contribute altogether to its achievement. Interactions between components and subsystems are usually classified into three main categories [1, 2]. Economic dependencies are involved when it is more advantageous (in terms of cost) to intervene on several components simultaneously rather than maintaining them after each other. The group interventions can afford to make savings in maintenance interventions (it moves a small number of times), or at the downtime of the system (when it shuts down the system to operate on a component), etc. Structural dependencies refer to systems where it is impossible to maintain a component without having an impact on others. This is particularly the case when a component is inaccessible directly and it is required to disassemble or stop another one in order to intervene. Stochastic dependencies occur when the state of a component may affect the lifetime distribution of other ones or when several components are subjected to common cause failures. Although the problem of stochastic, economic, and structural dependencies have been widely studied for maintenance issues [1, 3–6], the challenges for modeling are still very important, because of the diversity of situations that are arising from the industry. We must actually consider these dependencies either to optimize intervention dates by grouping or because we must take into account certain constraints (e.g., in the case of structural dependencies). The dependencies will condition the optimal specific actions for each component and under which conditions it is advantageous to group or ungroup interventions. To conventional maintenance optimization choices at the component level, it is therefore added that of deciding which components it is necessary to maintain altogether. Many approaches are proposed to account for this type of decision problem in the optimization process of maintenance policies. A classification of these methods is presented in [7], where the approaches are distinguished from “component to the system” and from “system to the component.” In the first case, the dates of interventions are optimized for each component independently and then

grouping strategies are considered by changing these optimal dates [8]. In the second case, it is often proposed to replace all the components simultaneously.

Generally speaking, the major challenge of the optimization scheme consists in joining the stochastic models describing the components' behavior with the combinatorial problems associating to maintenance grouping. Two kinds of models have been proposed in the literature [2]. Stationary models providing a long term or infinite planning horizon can be used in case of stable situations. Dynamic models can be used to change the planning rules according to short-term information (e.g., varying deterioration of components, spare parts, maintenance constraints, etc.) using a rolling horizon approach [2, 9]. The later is widely used and the reader can refer to [10] for an overview. However, in the models proposed and the associated optimization algorithms are usually limited to a specific problems with two significant hypotheses: (i) maintenance duration is small and can be neglected, (ii) only one preventive maintenance action for each component can be performed within the rolling horizon. Face to these issues, a maintenance model has been proposed in [11] to remove these two assumptions. In addition, in this works, maintenance opportunities, which are defined as inactivity periods of the system (with restricted duration) that may occur randomly over time, are used to carry out several maintenance (preventive or/and corrective) actions. Nevertheless, in a such model some important issues related to limited maintenance resources are not investigated. It is however pointed out in [1, 12, 13] that, from a practical point of view, it is often impossible to perform all the desirable maintenance actions because maintenance resources, such as maintenance budget, maintenance teams are limited. In the other hand, industrial systems are often asked to serve a sequence of missions with specific availability levels or limited stoppage durations [12, 14–17]. It should be however noticed that when integrating these constraints in the maintenance model, finding an optimal maintenance planning leads to solve NP-hard problems. Furthermore, when only maintenance teams are limited, a question arising is how to optimally allocate maintenance operations to each maintenance team.

This chapter focuses on a global approach allowing to overcome the mentioned issues. In that way, we will, on one hand, highlight a modeling framework which can taking into account the economic dependence between components, the maintenance duration, the maintenance occurrence in the considered horizon as well as the maintenance constraints such as the restricted maintenance duration and the limited maintenance teams. The presented maintenance models can also be applied in a “dynamic context” defined as a change on maintenance constraints, the occurrence of maintenance opportunities, etc. On the other hand, we will highlight the calculation and optimization tools to find an optimal grouped maintenance planning and to update it in presence of a dynamic context which may randomly occur on-line. The chapter is organized as follows. [Section 2](#) gives the description of general assumptions, maintenance operations, and costs. Maintenance constraints and opportunities are also discussed. [Section 3](#) focuses on the presentation of a general grouping approach under limited maintenance resources and duration. The application of the grouping approach through an example is described in [Section 4](#). In [Section 5](#), we will show how to integrate dynamically some maintenance opportunities in the grouping maintenance

approach presented in [Section 3](#). Finally, the last section concludes the chapter with a discussion of topics for future research.

2 Failure modeling and maintenance problems

2.1 Failure modeling, maintenance operations, and costs

Consider a system consisting of N components in which a preventive or corrective maintenance action on one or more components needs a shutdown of the entire system. Failure behavior of a component has been widely investigated and different failure models have been proposed and successfully applied in different applications [18]. The failure models are mainly classified into two categories: time-independent failure models or constant failure rate and time-dependent failure models. The first kind of models based mainly on exponential law are quite simple to implement however its application is limited since it cannot take into account the aging phenomena of the component. The second kind of failure models in which the failure rate usually increases with time are widely used in industry [2, 18]. The most popular time-dependent failure model is Weibull distribution. In that way, in this chapter it is assumed that the failure behavior of component i ($i = 1, \dots, N$) is described by a Weibull distribution with scale parameter $\lambda_i > 0$, and shape parameter $\beta_i > 1$. The failure rate of component i is then

$$r_i(t) = \frac{\beta_i}{\lambda_i} \left(\frac{t}{\lambda_i} \right)^{\beta_i - 1}. \quad (1)$$

The model parameters (λ_i, β_i) can be estimated from historical data, see [19].

To avoid the failure occurrence of a component, a preventive maintenance action is usually performed. In the literature, two kinds of preventive maintenance actions have been introduced [2, 20–22]. Perfect actions (or replacement), which can restore the maintained component to as good as new, have been studied in various maintenance models. Imperfect maintenance which can restore the system state to somewhere between the state before maintenance and as good as new. The implementation of imperfect maintenance policies are complex, especially for multicomponent systems. They have been applied mainly for single component systems, see for example [12, 23–25]. In this chapter, only perfect preventive maintenance action is considered and the maintenance cost for each preventive action on component i , denoted C_i^p , can be divided in three parts $C_i^p = S + c_i^p + c_i^d$:

- specific preventive cost c_i^p depending on the characteristics of component i ;
- setup cost, denoted S , indicates the logistic cost (or preparation cost). The setup cost can be shared when performing together several maintenance activities due to the economic dependence, e.g., only one setup cost is needed for executing a group maintenance activities [10]; and
- an unavailability cost c_i^d due to the system downtime during maintenance (resulting, e.g., from the production loss due to maintenance). Let d_i denote the preventive duration of component i (so-called activity i) including setup time (disassembly and reassembly times)

and replacement time, the unavailability cost $c_i^d = d_i \cdot C^d$ with C^d is the unavailability cost rate of the system. This additional cost can be also shared when several maintenance actions are performed together.

It is also assumed that if a component fails, it is then repaired immediately to put the system into its working state. This corrective maintenance action can restore the failed component into the state just before the failure (as bad as old). Under this corrective policy, the corrective duration is relatively small regarding to maintenance planning horizon and can usually be neglected. In that way, when a corrective action is performed on component i , a corrective cost denoted C_i^c is incurred. Without loss of generality, it is supposed that C_i^c includes already the specific corrective cost and the setup cost.

2.2 Maintenance constraints and opportunities

2.2.1 Maintenance constraints

From a practical point of view, industrial systems are often asked to serve a set of missions with given availability levels or limited breaks durations [12, 14–17]. In that way, in this chapter it is assumed that the system has to operate with a sequence of U missions. Each mission i ($i = 1, \dots, U$) is represented by three parameters: t_b^i , t_e^i , and D_0^i :

- t_b^i, t_e^i ($t_b^i < t_e^i$) are respectively the beginning and ending date of mission i . It is reasonable to assume that $t_e^i = t_b^{i+1}$; and
- D_0^i ($0 < D_0^i \leq t_e^i - t_b^i$) indicates the total duration allowed for executing all possible preventive actions within mission i .

In the other hand, it is shown in a large number of works that maintenance resources such as maintenance budget and maintenance teams (repairmen) are often limited, see for instance [1, 12, 13, 26]. To take into account this consideration, we assume that to execute preventive activities, only m maintenance teams are considered and each maintenance team can take only one preventive maintenance action at a time. In addition, the number of repairmen and/or the total maintenance duration allowed for a given mission may be changed in time due to economical/technical reasons. These kind of situations, which are referred to as “dynamic contexts,” may impact the maintenance planning and/or cost and should be taken in consideration into the maintenance optimization process [9].

2.2.2 Opportunities for maintenance

In industry, the system may be out of service during certain periods for whatever reasons, e.g., due to a change on the production or/and commercial planning. These inactivity periods should be seen as an interesting opportunity to carry out several maintenance actions because the maintenance cost may be dramatically reduced. Moreover, an inactivity period may occur randomly with time [8, 27–29]. In this situation, the question remains is how to integrate dynamically these opportunities into maintenance models. Detailed discussions will be presented in Section 5. Readers may also find a similar discussion in [11].

2.3 Need of dynamic grouping

Maintenance grouping strategy is of interest regarding to two following reasons:

1. It is shown in the literature, see for example [1, 10, 11], that grouping can save maintenance costs thanks to the sharing of the setup cost and/or the unavailability cost when several components are preventively maintained together. However, it is important to note that maintenance grouping may lead to some penalty costs related to indirectly penalized
 - a reduction on the useful life of components if their maintenance date are advanced; and
 - an increasing of components failure probability which could lead to a system shutdown if their maintenance date are too late.
2. When the total maintenance duration allowed is limited, several maintenance activities have to be grouped together to reduce maintenance durations. The grouping is needed even it may lead to a higher maintenance cost in some cases. As an example, when the setup cost and unavailability cost rate are neglected, grouping becomes costly but without grouping the limited duration constraint may not be reached.

In addition, such a maintenance grouping should be able to efficiently updated in presence of a dynamic context which may randomly occur in time. The later will be discussed in detail in [Sections 4 and 5](#).

3 Grouping maintenance approach

The grouping maintenance approach is divided into four steps: individual maintenance optimization, tentative planning, grouping optimization, and updating. The illustration of this approach is shown in [Fig. 1](#).

3.1 Step 1: Individual maintenance optimization

The objective of this step is to find the optimal preventive maintenance period for each component. To do that, an infinite-horizon maintenance model can be used by considering an average use of component i . The interactions between components are neglected in this step.

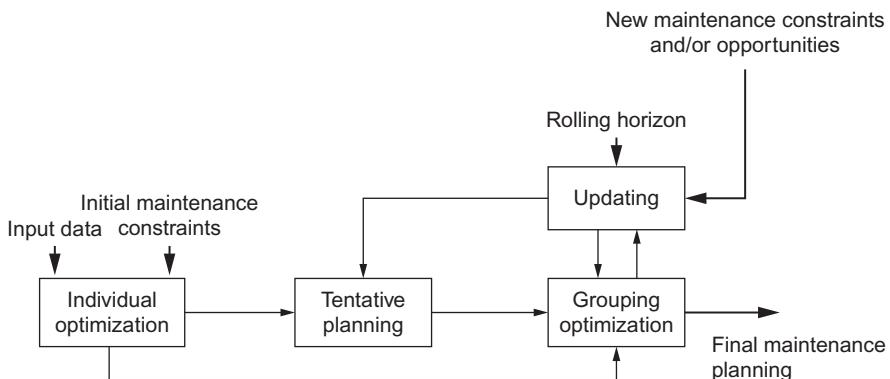


Fig. 1 Maintenance grouping approach.

From a mathematical point of view, the expected deterioration cost for component i within the period $(0, x]$, denoted $M_i(x)$, can be expressed as:

$$M_i(x) = C_i^c \cdot \int_0^x r_i(y) dy. \quad (2)$$

From Eqs. (1), (2), we obtain:

$$M_i(x) = C_i^c \cdot \left(\frac{x}{\lambda_i} \right)^{\beta_i}. \quad (3)$$

If a preventive maintenance action on component i is performed at time x , the expected cost within the interval $[0, x + d_i]$ is then:

$$\Gamma_i(x) = C_i^p + M_i(x) = C_i^p + C_i^c \cdot \left(\frac{x}{\lambda_i} \right)^{\beta_i}. \quad (4)$$

According to the renewal theory [19], if component i is preventively maintained every x time units, the average maintenance cost per time unit, denoted $\phi_i(x)$, is then calculated by:

$$\phi_i(x) = \frac{\Gamma_i(x)}{x} = \frac{C_i^p + C_i^c \cdot \left(\frac{x}{\lambda_i} \right)^{\beta_i}}{x}. \quad (5)$$

The optimal maintenance period, x_i^* , is given when the average maintenance cost rate reaches its minimal value $\phi_i^* = \phi_i(x_i^*)$. From Eq. (5), we obtain:

$$x_i^* = \lambda_i^{\beta_i} \sqrt{\frac{C_i^p}{C_i^c(\beta_i - 1)}}, \quad (6)$$

and the minimum average cost rate:

$$\phi_i^* = \phi_i(x_i^*) = \frac{C_i^p \beta_i}{x_i^*(\beta_i - 1)}. \quad (7)$$

If all components of the system are individually maintained, the minimum total maintenance cost per operating time unit of the system is calculated as follows:

$$C_{IM} = \sum_{i=0}^n \phi_i^*. \quad (8)$$

It should be noticed that x_i^* representing the nominal preventive maintenance period of component i ($i = 1, \dots, N$) will be used to establish a tentative maintenance planning.

3.2 Step 2: Tentative planning

To establish a tentative planning, a planning horizon denoted $[t_{begin}, t_{end}]$ need to be firstly defined. Indeed, t_{begin} corresponds to the current date and the planning horizon should be chosen in the way so that all components are concerned in the maintenance decision. The later means that chosen horizon ensures that each component is preventively maintained at least one time.

Based on the nominal preventive maintenance periods given in the previous step, the first tentative maintenance execution time of each component i ($i = 1, \dots, n$) within the considered horizon is expressed as:

$$t_{i1} = t_{begin} - t_i^e + x_i^* + d_i^\Sigma, \quad (9)$$

with

- t_i^e is the total operating time of component i elapsed from its last replacement before t_{begin} ; and
- d_i^Σ indicates the cumulative maintenance durations before the tentative maintenance date of component i .

The illustration of t_{i1} is presented in Fig. 2.

Since the nominal preventive period of components may be different, several components could be preventively replaced more than one time within the considered horizon. Let i^j be the j th maintenance occurrence of component i (activity i), the tentative maintenance date of operation i^j , denoted t_{ij} ($j \geq 2$), depends on the executed date of operation i^{j-1} (denoted t_{ij-1}^*), the cumulative maintenance durations d_{ij-1}^Σ from t_{ij-1}^* as well as the nominal maintenance period x_i^* .

$$t_{ij} = t_{ij-1}^* + d_{ij-1}^\Sigma + x_i^*. \quad (10)$$

Finally, t_{end} can be determined as follows:

$$t_{end} = \begin{cases} t_e^U & \text{if } t_e^U > (t_{j1} + d_j) \text{ with } t_{j1} = \max_{i=1:n} t_{i1}; \\ t_{j1} + d_j & \text{otherwise.} \end{cases} \quad (11)$$

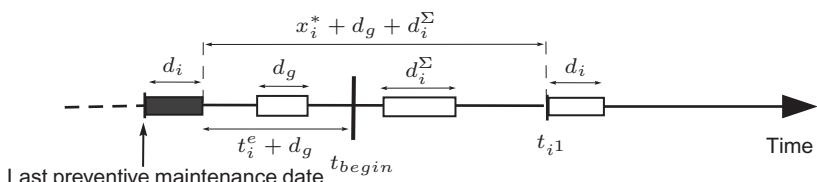


Fig. 2 Illustration of t_{i1} .

3.3 Step 3: Maintenance grouping optimization

The objective of this step is to find, at system level, an optimal maintenance planning that minimizes the total maintenance cost and meets the maintenance constraints (limited maintenance duration and maintenance teams). The main idea is herein to find a grouping structure (or a partition of all maintenance activities/operations in the considered horizon) in which at each maintenance date, several maintenance activities are jointly performed, so-called grouped maintenance actions. In that way, each group is characterized by the maintenance operations, the execution date and the group maintenance duration. This step is divided into two phases.

3.3.1 Phase 1: Mathematical formulation

Suppose that several different preventive maintenance activities $i^j (i, j = 1, 2, \dots)$ are carried out together in a group G^k (with $k = 1, 2, \dots$). It should be noticed that operations i^j and $i^{j'}$ ($j \neq j'$) are identical since they are respectively the j th and the j' th maintenance occurrence of component i . As a consequence, i^j and $i^{j'}$ cannot be together in a same group.

The total economic profit (or saving cost) of a given group G^k can be divided into different parts as follow:

1. The setup cost savings: let S_{G^k} be the setup cost for executing the maintenance group G^k , the setup cost saving when compared with individual maintenance is:

$$V_1(G^k) = \text{card}(G^k) \cdot S - S_{G^k}, \quad (12)$$

where $\text{card}(G^k)$ is the number of components in group G^k . According to [10, 11], executing a group components requires only one setup cost, in that way $V_1(G^k)$ can be calculated as follows:

$$V_1(G^k) = (\text{card}(G^k) - 1) \cdot S. \quad (13)$$

It is clear that $V_1(G^k)$ depends on the number of group components. Note also that other models for setup cost sharing can be found in [30].

2. An additional cost saving associating with the reduction on maintenance duration due to maintenance grouping. The duration reduction depends not only on the characteristics of the group but also on the number of maintenance teams. Indeed, this cost saving can be written as:

$$V_2(G^k, m) = \left(\sum_{i^j \in G^k} d_i - d_{G^k}(m) \right) C^d, \quad (14)$$

where $d_{G^k}(m)$ is the total duration of group G^k . $d_{G^k}(m)$ depends on both the group components' maintenance duration and the number of maintenance teams m . If only one team is available, then $d_{G^k}(1) = \sum_{i^j \in G^k} d_i$, as a consequence $V_2(G^k, m) = 0$. For $m > 1$, in order to optimally allocate the maintenance activities for each maintenance team and to find the minimum value of $d_{G^k}(m)$, an optimization algorithm, so called MULTIFIT algorithm, can be used. For the detailed implementation of this algorithm, the reader may refer to [26].

3. As pointed earlier, grouping may lead also to a penalty costs due to the changes of the nominal maintenance period. Let $h_i(\Delta t_{ij})$ be the penalty cost when t_{ij}^* is actually executed at time $t_{ij}^* = t_{ij} + \Delta t_{ij}$ (with $\Delta t_{ij} > -x_i^*$) instead of t_{ij} . It is pointed out in a large number of works, see for instance [1, 10, 11], that the change on the maintenance execution date may lead to:

- an increase in the expected costs of the j th renewal cycle, that is given by $M_i(x_i^* + \Delta t_{ij}) - M_i(x_i^*)$; and
- a change cost relying on the deferments of all future executions maintenance dates after t_{ij} that are given by $\Delta t_{ij} \cdot \phi_i^*$.

In that way, $h_i(\Delta t_{ij})$ can be calculated as:

$$h_i(\Delta t_{ij}) = C_i^c \cdot \left[\left(\frac{x_i^* + \Delta t_{ij}}{\lambda_i} \right)^{\beta_i} - \left(\frac{x_i^*}{\lambda_i} \right)^{\beta_i} \right] - \Delta t_{ij} \frac{C_i^p \beta_i}{x_i^* (\beta_i - 1)}. \quad (15)$$

It should be noticed that another model for penalty cost function can be found in [31].

Let $H_{G^k}(t)$ be the penalty cost function of the group G^k at time t . The optimal execution date of the group, denoted t_{G^k} , can be given when the $H_{G^k}(\cdot)$ reached its minimum value $H_{G^k}^*$. That is:

$$H_{G^k}^* = H_{G^k}(t_{G^k}) = \min_t \left(\sum_{i \in G^k} h_i(t - t_{ij}) \right). \quad (16)$$

t_{G^k} can be numerically determined.

According to Eqs. (16), (13), (14), the cost saving of group G^k , namely Q_{G^k} , can be calculated as follows:

$$Q_{G^k} = V_1(G^k) + V_2(G^k, m) - H_{G^k}^*. \quad (17)$$

From all different groups, a grouping structure, namely GS , can be established. In fact, GS is a collection of s mutually exclusive groups G^1, \dots, G^s which cover all preventive maintenance activities in the considered horizon, i.e.,

$$G^l \cap G^k = \emptyset, \quad \forall l \neq k \text{ and } G^1 \cup G^2 \cup \dots \cup G^s = \{1, \dots, N\}. \quad (18)$$

In that way, we obtain the total economic profit (or saving cost) of grouping structure GS as follows:

$$Q_{GS} = \sum_{G^k \in GS} Q_{G^k}. \quad (19)$$

The total preventive duration for a given mission j ($j = 1, \dots, U$) can be determined by:

$$D^j = \sum_{t_b^j \leq t_{G^k} < t_e^j} d_{G^k}(m). \quad (20)$$

According to the limited maintenance duration constraint presented in [Section 2.2](#), a grouping structure is an optimal one, denoted GS^* , if it satisfies the following conditions:

$$GS^* = \underset{GS}{\operatorname{argmax}} Q_{GS}, \quad (21)$$

and,

$$D^i \leq D_0^i \text{ for } i = 1, \dots, U. \quad (22)$$

After identifying the optimal grouping structure, the minimum maintenance cost rate for the considered planning horizon can be evaluated by:

$$C_{grouping} = C_{IM} - Q_{GS^*} / (t_{end} - t_{begin} - D_{GS^*}). \quad (23)$$

3.3.2 Phase 2: Finding an optimal grouping structure

Due to the impacts of limited maintenance duration constraints implying the consideration of any combinations of all possible groups, the identifying of an optimal grouping structure (or grouped maintenance planning) leads to a NP-complete problem to be solved. Indeed, the number of possible grouping structures rapidly increases regarding to an increasing on the number of components. As a consequence, analytical methods are computationally burdensome and almost unusable or highly inefficient. It is pointed out in [\[9\]](#) that analytical methods can be usable when the number of components is lower than 14.

It should be noticed that in the case of one maintenance team without limited maintenance duration constraint, to reduce the computing time a theorem, so-called the theorem of consecutive preventive activities, has been introduced in [\[10\]](#). However, it is shown in [\[26\]](#) that this theorem is no longer applicable in presence of limited maintenance teams or/and restricted duration constraint. To face this issue, Generic Algorithms (GA) have been successfully applied for finding an optimal grouped maintenance planning. For the detailed implementation of GA algorithm, the reader may refer to [\[26\]](#).

3.4 Step 4: Updating of the grouped maintenance planning

The current optimal maintenance planning given by the previous step can be used in stable situation but it needs to be updated in presence of a dynamic context such as in presence of new maintenance constraints, occurrence of an maintenance opportunity or at the end of the current horizon. More precisely:

- 1. in presence of new maintenance constraints.** From a practical point of view, maintenance constraints may be changed in time, e.g., several maintenance teams may become unavailable during given time periods and/or the total duration allowed for maintenance

of given missions may be changed. These changes may directly impact the current maintenance planning, i.e., the current planning may be no longer an optimal one or even become unusable one. As a consequence, in these cases, the current maintenance planning should be updated by taking in consideration of new maintenance constraints. To do that, we may go back to step 2 in order to redefine all preventive maintenance activities in the new tentative planning horizon and so on. An illustration on the updating process will be presented in [Section 4](#). It should also be noticed that updating the grouped maintenance planning may be also required with other kinds of dynamic contexts such as change of production planning, spares parts replenishment, etc. [9].

2. ***in presence of maintenance opportunities.*** As pointed in [Section 2](#), an inactivity period or maintenance opportunity may become occur in time. In presence of an opportunity, the current grouped maintenance planning should be updated taking in consideration of this opportunity. The later will be presented in [Section 5](#).
3. ***at the end of the current planning horizon.*** A new maintenance planning for the next horizon interval is required. To this end, new horizon need first to be defined regarding to new required missions' profile (time intervals, maintenance duration allowed and/or available maintenance teams). The rolling horizon procedure is then applied by returning to step 2 and so on.

4 Numerical study

Numerical results are presented in this section for a system of 20 series units. The aim is to demonstrate how the aforementioned approach is applied with limited maintenance duration, and limited repair team. When a unit fails, a minimal-repair (as bad as old) is performed immediately. Preventive replacements are performed with nonnegligible maintenance duration. The lifetime distribution of unit i ($i = 1, \dots, 20$) is a Weibull one (scale parameter $\lambda_i > 0$, shape parameter $\beta_i > 1$). The lifetime parameters and maintenance costs are given in [Table 1](#) (arbitrary time unit (atu) and arbitrary cost unit (acu) are given, e.g., in [Table 1](#) λ_i, d_i, t_i^e are given in atu and c_i^p, C_i^c in acu). The setup cost and unavailability cost rate are respectively $S = 10$ (acu) and $C^d = 5$ (acu/atu). Regarding individual planning, optimal maintenance periods x_i^* , minimum maintenance cost rates ϕ_i^* , and the next preventive replacement date t_{i^1} (with $i = 1, \dots, 20$) are obtained with Eqs. (6), (9). [Table 2](#) gives the results. Hence if individual preventive maintenance actions are performed independently, the minimum maintenance cost is:

$$C_{IM} = \sum_{i=1}^{20} \phi_i^* = 19.75.$$

Here, the planning horizon is from 0 ($t_{begin} = 0$) to $t_{20^1} + d_{20} = 605$ ($t_{end} = 605$). The cumulated duration of the maintenance equals $\sum_{i=1}^{20} d_i = 71$ and the system availability is in average: $(t_{end} - t_{begin} - \sum_{i=1}^{20} d_i) / (t_{end} - t_{begin}) = 0.8826$. In the planning horizon, the corresponding maintenance cost is:

$$TC_{IM} = C_{IM} \cdot (605 - 71) = 10,546.5.$$

Table 1 Parameters for a 20-component system

Component	λ_i	β_i	c_i^p	C_i^c	d_i	t_i^e	Component	λ_i	β_i	c_i^p	C_i^c	d_i	t_i^e
1	237	1.5155	266	79	1	847.7	11	209	1.8171	380	52	2	461.9
2	255	1.3981	347	67	2	1614.1	12	297	1.5439	376	56	5	1327.3
3	335	1.6527	322	94	6	903.6	13	326	2.0178	279	80	5	326.4
4	291	1.8663	362	85	2	602.1	14	236	1.5127	249	59	1	663.3
5	186	1.3280	500	59	3	2122.1	15	278	1.2336	221	66	5	2357.4
6	263	2.0819	401	60	4	466.9	16	169	1.7985	342	85	2	69.7
7	260	2.1427	326	100	3	247.0	17	281	1.6067	446	86	6	750.0
8	243	2.1816	247	43	5	324.1	18	257	1.5251	232	81	4	405.6
9	226	1.3909	329	75	4	1144.7	19	235	1.2908	280	58	3	1726.6
10	268	1.4262	316	80	3	1096.9	20	239	1.7271	338	24	5	873.4

Table 2 Values of x_i^* , ϕ_i^* , and t_{i^l}

Component	x_i^*	ϕ_i^*	t_{i^l}	Component	x_i^*	ϕ_i^*	t_{i^l}
1	847.7	0.9745	0	11	717.9	1.2391	289
2	1663.1	0.7750	50	12	1602.3	0.7281	310
3	980.6	0.9348	80	13	636.4	0.9782	350
4	703.1	1.1705	110	14	988.3	0.7881	370
5	2231.1	0.9519	120	15	2711.4	0.4986	400
6	652.9	1.2703	200	16	428.7	1.9021	410
7	439.0	1.4994	210	17	1127.0	1.1421	430
8	533.1	0.9767	230	18	846.6	0.8989	500
9	1368.7	0.9333	250	19	2213.6	0.6116	550
10	1364.9	0.8472	280	20	1407.4	0.6295	600

The added value of the grouping strategy is hereafter demonstrated in the following cases:

- with no constraint;
- with limited maintenance duration; and
- with limited maintenance duration and limited repairmen.

4.1 No constraint

The mission time of the system is supposed to be $[0, 605]$. [Table 3](#) gives the optimal grouping strategy with no constraint: it is made of two optimal groups $G^1 = \{1, \dots, 11\}$ and $G^2 = \{12, \dots, 20\}$.

To exemplify, [Fig. 3](#) shows the optimal solution for G^2 with seven repairmen (MULTIFIT algorithm is used).

[Table 3](#) shows the maintenance cost saving: $219.66 + 219.32 = 438.98$ (around 3.75% of TC_{IM}). In addition, it is worth to note that the maintenance duration is significantly decreased, $D_{GS^*} = 12$. Conversely the average system availability increases a lot: $(605 - 12)/605 = 0.9802$.

Table 3 Grouping without constraints

Group components	Optimal date	Duration	Saving
{1,...,11}	173.3	6	219.66
{12,...,20}	364.8	6	219.32

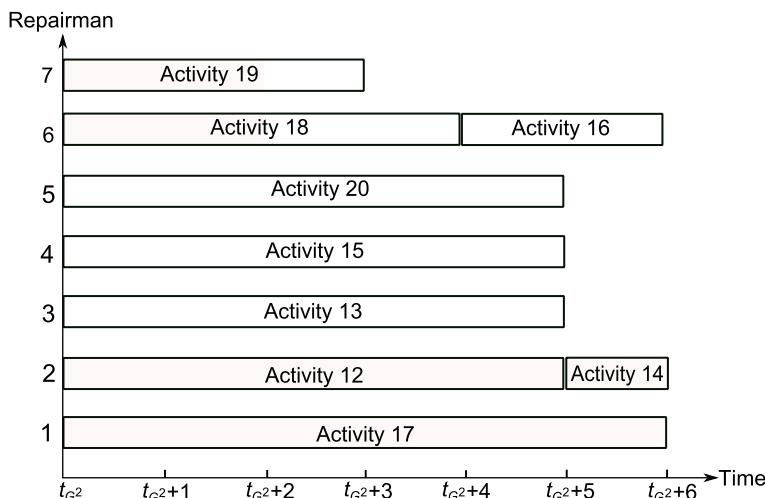


Fig. 3 Optimal solution for $G^2 = \{12, \dots, 20\}$ with seven repairmen.

The optimal solution indicates also that seven repairmen can implement the maintenance planning. It can be useful input for the maintenance manager.

4.2 Limited repairmen number

The number of repairmen m can vary from 1 to 10 without additional cost and an optimal solution is given for each m (see Table 4). The structure of the grouping strategy is highly impacted by the value of m .

For each optimal grouping strategy, the average system availability over the planning horizon is given by:

$$A = \frac{t_{end} - t_{begin} - D_{GS^*}}{t_{end} - t_{begin}}. \quad (24)$$

[Fig. 4](#) illustrates the effects of the repairmen number.

All the results show that if the repairmen are less than 7 ($m \leq 7$), their number can really impact the maintenance cost and the average availability. Actually, the maintenance duration can in this case be significantly decreased by adding repairmen. Conversely, when $m \geq 7$, the repairmen number has no more effect neither on the maintenance cost, nor on the average availability.

Table 4 Grouped maintenance planning under limited repairmen

m	Optimal solution	t_G	d_G	D_{GS^*}	Q_{GS^*}
1	$G^1 = \{1, \dots, 5\}$	71.3	14	71	154.51
	$G^2 = \{6, \dots, 12\}$	218.9	26		
	$G^3 = \{13, \dots, 20\}$	401.7	31		
2	$G1 = \{1, \dots, 5\}$	71.3	7	36	329.51
	$G2 = \{6, \dots, 12\}$	211.9	13		
	$G3 = \{13, \dots, 20\}$	381.7	16		
3	$G^1 = \{1, \dots, 5, 9\}$	83.9	6	24	386.32
	$G^2 = \{6, 7, 8, 10, 11\}$	208.8	6		
	$G^3 = \{12, \dots, 20\}$	370.8	12		
4	$G^1 = \{1, \dots, 5, 9, 15\}$	87.4	6	19	410.19
	$G^2 = \{6, 7, 8, 10, 11, 12\}$	210.4	6		
	$G^3 = \{13, 14, 16, \dots, 20\}$	373.8	7		
5	$G^1 = \{1, 2, 4\}$	68.9	2	16	423.41
	$G^2 = \{3, 5, \dots, 12, 15\}$	199.0	8		
	$G^3 = \{13, 14, 16, \dots, 20\}$	371.8	6		
6	$G1 = \{1, \dots, 11\}$	173.3	6	13	433.98
	$G^2 = \{12, \dots, 20\}$	364.8	7		
≥ 7	$G^1 = \{1, \dots, 11\}$	173.3	6	12	438.98
	$G^2 = \{12, \dots, 20\}$	364.8	6		

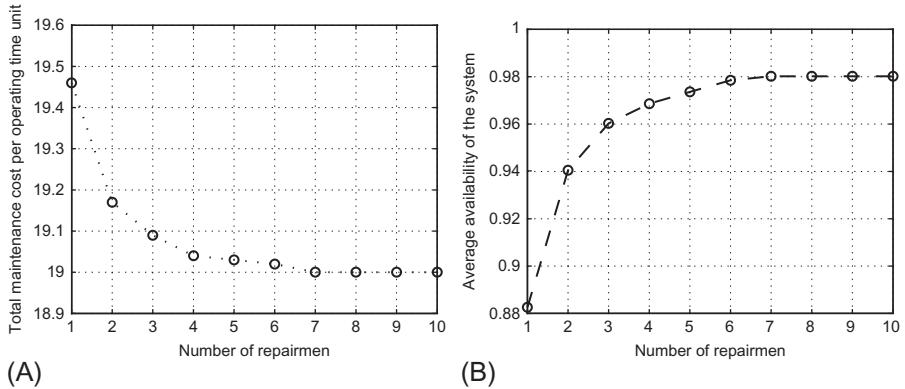


Fig. 4 Impact of the limited repairmen on (A) the total maintenance cost rate and (B) the average availability A .

4.3 Limited maintenance duration

The mission time of the system is still supposed to be $[0, 605]$. A limited duration (denoted D_0) is considered for the maintenance interventions and the obtained maintenance cost is given in Fig. 5 for different values of D_0 .

When $D_0 \leq 12$, the maintenance cost increases if the maintenance duration decreases. This is all the more true when $D_0 \leq 3$. It is worth to note that a given maintenance duration constraint can be achieved by increasing the repairmen number. This is illustrated in Fig. 5B. If $D_0 \geq 11$, seven repairmen are required to perform the optimal planning. If $D_0 \leq 7$, the required repairmen number is greater with an optimal at 13.

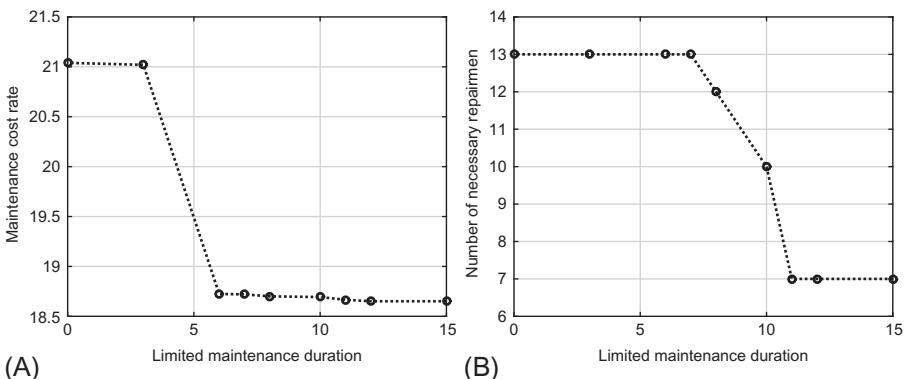


Fig. 5 Impact of required average availability level on (A) the total maintenance cost rate and (B) the number of necessary repairmen.

4.4 Limited maintenance duration and limited repairmen number

We consider now that both the time to repair and the number of people to operate are limited. The same mission time is still considered. Different values of D_0 are investigated and for each, the repairmen number varies from 1 to 15. Results are given in **Table 5**. The numerical results demonstrate:

- the limits of the maintenance planning when strong constraints have to be taken into account. Repairmen number has to be optimized for each value of D_0 . The lower bound is the minimum number of repairmen required to perform maintenance planning. The upper bound is the number of repairmen that corresponds to the lowest maintenance cost. As an illustration, let us look at: $D_0 = 10$ (average availability has to be higher than 0.9835). If we have less than eight repairmen ($m < 8$), it is impossible to establish any maintenance planning. Then if $m \geq 10$, there is no more variations of the maintenance cost. It is worth to note that the upper bound of repairmen is lower than the maximum number of maintenance activities in a group (e.g., for $D_0 = 7$ 13 repairmen are required to perform maintenance actions for a group of $19 G^2 = \{2, 3, \dots, 20\}$);
- the impact of the repairmen capacity when the number of repairmen is in the bounded interval.

5 Dynamic grouping and maintenance opportunities

The aforementioned grouping strategy is now used in a dynamic context when maintenance opportunities arise randomly (typically when the whole system is stopped).

5.1 Maintenance opportunities modeling

An opportunity v ($v = 1, 2, \dots$) is defined with three parameters t , T_{opp}^v , and D_{opp}^v :

- t is the time upon which a new opportunity v is known to be existing;
- T_{opp}^v is the starting time of v ; and
- D_{opp}^v is the time length of v .

The system is supposed to be unavailable during the whole opportunity length. This implies that there is no unavailability cost c_i^d related to production loss if the maintenance is performed at this time. The cost involved is reduced to $S + c_i^p$. The aim is then to re-schedule on-line maintenance interventions to take advantage of this unexpected cost saving.

5.2 Updating of maintenance planning

An optimal grouped maintenance planning is first calculated in the time window $[t_{begin}, t_{end}]$. At t ($t_{begin} < t \leq t_{end}$) an opportunity is known to occur in the future at T_{opp}^v ($T_{opp}^v > t$) with time length D_{opp}^v . The optimal maintenance planning is then updated. Note that if at t , a group is being currently maintained, the updated

Table 5 Optimal maintenance planning with limited maintenance duration and limited repairmen number

Limited maintenance duration D_0	Repairmen constraint	Optimal solution	t_G	d_G	Q_{GS^*}	D_{GS^*}	Necessary repairmen m^*
7	$m < 11$	No solution					
	$m = 11$ or $m = 12$	$G^1 = \{1, \dots, 20\}$	241.45	7	391.74	7	11
	$m \geq 13$	$G^1 = \{1\}$ $G^2 = \{2, \dots, 20\}$	0 249.06	1 6	398.76	7	13
10	$m < 8$	No solution					
	$m = 8$	$G^1 = \{1\}$ $G^2 = \{2, \dots, 20\}$	0 249.06	1 9	383.76	10	8
	$m = 9$	$G^1 = \{1, 2, 4, 5, 6, 7, 9, 19\}$ $G^2 = \{3, 8, 10, \dots, 18, 20\}$	159.82 304.27	4 6	416.50	10	9
	$m \geq 10$	$G^1 = \{1, 2, 4, 5, 6, 7, 9\}$ $G^2 = \{3, 8, 10, \dots, 20\}$	156.58 305.55	4 6	419.57	10	10

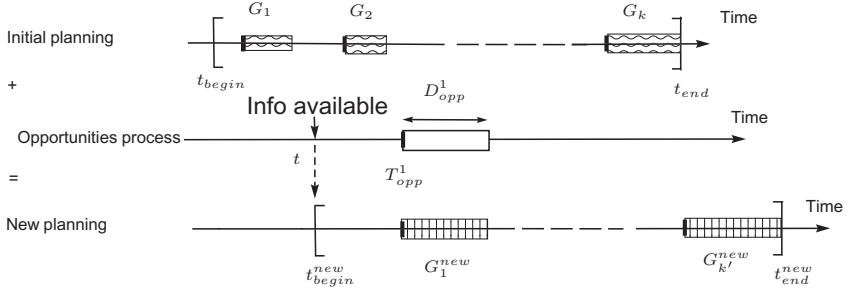


Fig. 6 Illustration of maintenance planning update with opportunity.

maintenance schedule will be applied after completion of the current task. [Fig. 6](#) gives an illustration of the updating scheme.

New groups are optimized by going back to step 2 (i) to update the scheduling horizon such that all units can be candidate for maintenance actions, and (ii) to calculate the first tentative execution times within this new horizon. Then new groups are optimized in step 3, according to the algorithm given in [11]:

1. Add a fictive maintenance operation opp , with tentative execution time T_{opp}^v and duration $d_{opp} = 0$. If opp is in an admissible group, the group must be then maintained at T_{opp}^v .
2. A group $G^p = \{..., opp, ..., p\}$ with $1 \leq opp \leq p$, is said to be admissible if:
 - T_{opp}^v is the optimal maintenance date;
 - $\sum_{i \in G^p} d_i \leq D_{opp}^v$ (total operations duration of the group is lower or equal to the opportunity length);
 - the maintenance operations in G^p are not the same;
 - G^p is irreducible (the group cannot be divided into two or more subgroups that give a better solution).

Then, if G^p is an admissible group, Eq. (14) gives:

$$V_2(G^p, m) = \sum_{i \in G^p} d_i C^d. \quad (25)$$

5.3 Numerical study

We focus here on a five units series system. After a failure, a unit is immediately and minimally repaired (as bad as old). The lifetime of each unit i ($i = 1, \dots, 5$) follows a Weibull distribution (scale parameter $\lambda_i > 0$, shape parameter $\beta_i > 1$). Data for the units are given in [Table 6](#). We take $S = 30$ and $C^d = 10$. We consider one mission horizon without time constraint and with only one repairman. It is then possible to use dynamic programming instead of GA algorithm to find the optimal grouped maintenance planning (readers can refer to [10, 32] for more detail).

C_i^p , x_i^* , ϕ_i^* , and t_{i^1} are deduced by substitution with Eqs. (6), (7), (9), see [Table 7](#).

Table 6 Data for a five-component system

Component i	1	2	3	4	5
λ_i	159	108	49	97	84
β_i	2.7	1.7	1.25	1.75	1.5
c_i^p	225	585	105	345	345
C_i^c	182	172	150	50	100
d_i	5	3	4	2	8
t_i^e	108	159	137	130	306

Table 7 Values of x_i^* , ϕ_i^* , and t_{i1}

Component i	1	2	3	4	5
C_i^p	275	615	145	364	425
x_i^*	158.2	289.9	168	372.5	366.1
ϕ_i^*	3	5.4	5.2	2.5	3.7
t_{i1}	54.2	147.9	30	271.5	69.1

5.3.1 Grouped maintenance planning

The scheduling horizon is defined with Eq. (11): $t_{end} = t_{41} + d_4 = 273.5$ units of time.

The optimal individual planning x_i^* ($i = 1, \dots, 5$) is shown in Fig. 7. Preventive maintenance durations are represented by red (dark gray in print versions) segments. Units 1 and 3 are supposed to be maintained two times while the others are supposed to be maintained once within $[0, t_{end}]$. The total downtime of the system is $D^\Sigma = 2.d^1 + d^2 + 2.d^3 + d^4 + d^5 = 31$. The maintenance costs expected for this horizon is:

$$TC1 = \sum_{i=1}^5 (t_{end} - D^\Sigma) \cdot \phi_i^* = 4819.45.$$

The grouping strategy is then given in Table 8: $G^1 = \{5, 6, 7\}$, $G^2 = \{4\}$, and $G^3 = \{1, 2, 3\}$. The total subsequent cost saving is $Q_\Sigma^* = 57.36 + 0 + 56.84 = 114.2$ which corresponds to 2.37% compared to individual maintenance actions cost.

The grouped maintenance planning is given in Fig. 8.

5.3.2 On-line updating of maintenance planning

We now consider opportunities arising on-line. We suppose that at $t = 100$ an opportunity is known to occur later at $T_{opp}^1 = 100$ with a limited length $D_{opp}^1 = 15$. The maintenance planning is updated by following the proposed approach and by going back to step 2. A new tentative planning is then defined with the new scheduling horizon

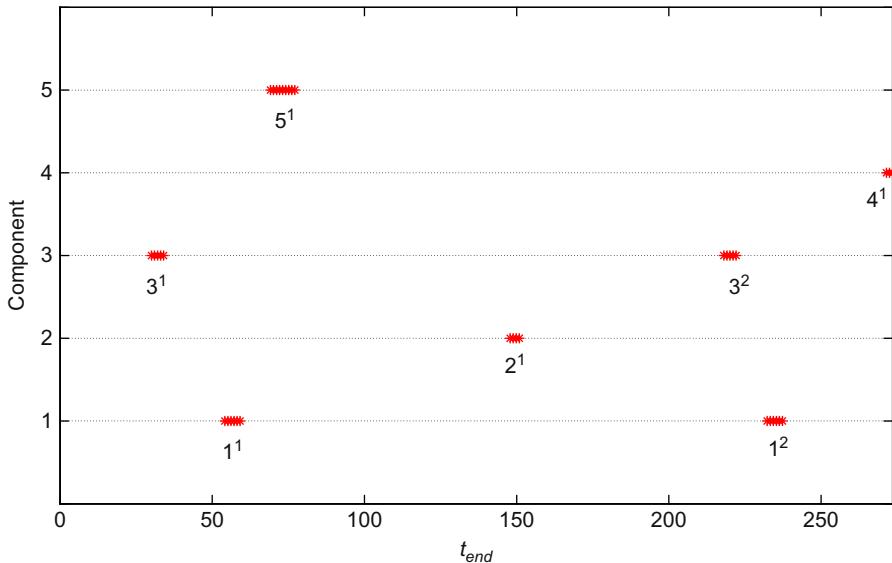


Fig. 7 Maintenance planning with individual maintenance optimization.

Table 8 Optimal groups and saving cost

Optimal group	Optimal date	Duration	Savings
$G^1 = \{1, 2, 3\}$	51.9	17	57.36
$G^2 = \{4\}$	147.9	3	0
$G^3 = \{5, 6, 7\}$	240	11	56.84

[100, 486]. The algorithm presented in [Section 5.2](#) is applied. A fictive maintenance operation (operation number 2) is added with tentative execution time $t_{1^1} = T_{opp}^1 = 100$ and duration $d^2 = 0$. The optimal groups obtained are: $G^1 = \{1, 2, 3, 4, 5\}$ and $G^2 = \{6, 7, 8\}$, see [Table 9](#). The total cost saving compared to individual replacements in the same horizon is $Q_{\Sigma}^* = 37.1 + 146.96 = 184.06$.

New groups after updating are given in [Fig. 9](#). Opportunity (called cost-effective one) is used for units 1, 2, 3, 4 to perform preventive replacements. It is worth to note that the maintenance cost gain will depend on t , T_{opp}^1 , and D_{opp}^1 .

6 Summary and conclusions

The present paper gives a complete modeling framework to deal with maintenance optimization problems at the strategic and at the operational levels. Strategic and operational levels are usually referring to different kinds of decision-making problems in

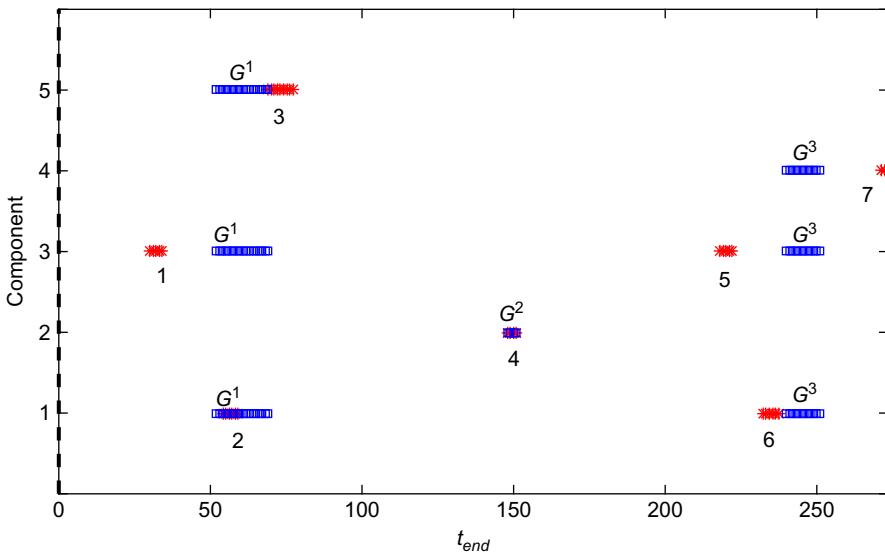


Fig. 8 Grouped maintenance planning.

Table 9 Grouping structure updated in presence of an opportunity

Optimal group	Optimal date	Duration	Savings
$G^1 = \{1, 2, 3, 4, 5\}$	200	14	146.96
$G^2 = \{6, 7, 8\}$	387.1	17	37.1

the context of maintenance management. In the first case, the decision has to be made in the long term, based on past experience for similar systems and situations. In the second case, the decision has to be made in the short term, based on on-line information available for the system under consideration. Optimal decisions for each level rely on different modeling methods and require to address several challenges. We intend to provide a unified approach to take into consideration both levels and to propose new solutions in this context.

From the strategic point of view, the challenge is to plan preventive maintenance activities under the assumption that the operational conditions are stationary, and that a long term horizon is meaningful. In this case, the maintenance plan is usually optimized in two steps, one for individual planning, one for grouping interventions. Typical influent parameters for the optimal solution are failure rates of the individual units into the system and setup costs. This optimization problem is well known and solved with exact methods under very restrictive hypotheses: series structure, only one repairman available, neglected maintenance duration, and no constraints. Such hypotheses guaranty that the optimal solution will be made of groups with units that have the most similar individual plannings [10, 11]. This theorem allows to reduce the

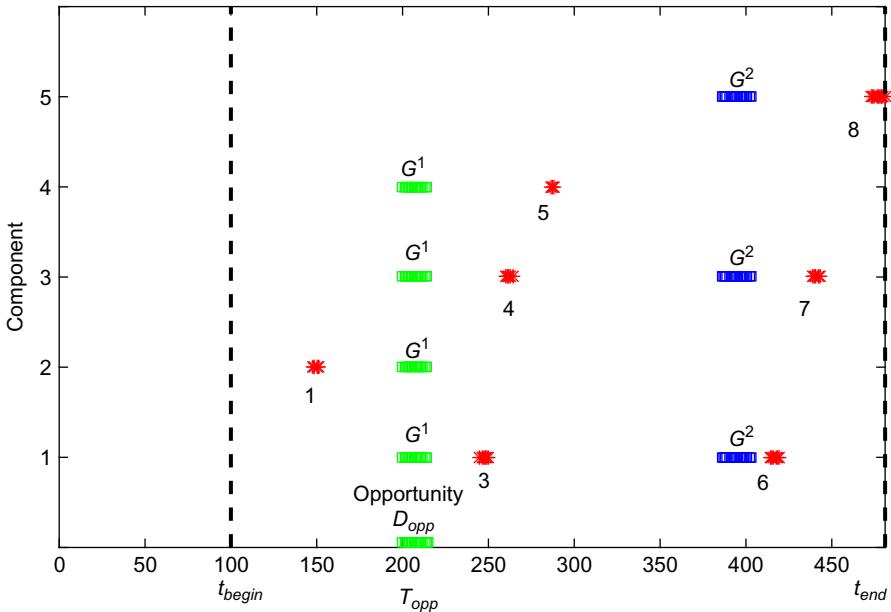


Fig. 9 Grouped maintenance planning updated with an opportunity.

number of possible solutions to explore. During the past ten years some papers [21] proposed to relieve one of the hypotheses and to use empirical methods (mainly based on genetic algorithms) to solve the optimization problem. The present chapter proposes a more generic framework to relieve most of the restrictive hypotheses: only the series structure assumption is kept.

From the operational point of view, the challenge is to take advantage of arising opportunities to perform preventive maintenance. It is typically the case when the system is put out of service for external reasons linked to the production and to the commercial planning, or when another upstream system fails. These inactivity periods may occur randomly and should be considered as interesting opportunities for maintenance actions. Then the main idea is to re-use dynamically the optimization algorithm developed for the strategic level. It is worth to note that in the current numerical results, we assume that there is no constraint and there is only one repairman. Hence the dynamic optimization is led with exact solution instead of genetic algorithm.

To exemplify more precisely possible outputs, we can say that the proposed modeling approach help to: (i) provide the minimum number of repairmen to ensure that an establishable maintenance planning, which copes with given restricted maintenance duration constraint, can be constructed; (ii) determine the minimum number of repairmen that leads to an optimal maintenance planning satisfying the restricted maintenance duration/availability constraints with lowest maintenance cost; (iii) update

optimally the maintenance planning in presence of new maintenance constraints and/or some opportunities which may occur randomly over time.

The presented maintenance approach focuses mainly on basic systems with fixed series structure (a single failure results in system downtime) and on economic dependencies. For more complex structure systems, readers can refer to some recent works, see for instance [9, 33]. But the whole modeling framework needs still to be strengthened to address challenges related to modularization and reconfigurations. In the recent works, complex structure concept is reduced to combinations of parallel and series units. Nothing is ready for more advanced complexity related to structural and stochastic dependencies between units. For system that are difficult to access and to maintain, it is often the case that some subparts of the system need to be removed as a whole (modularization of subsea compressors for example). Then the grouping solutions can be completely affected by this constraints at the strategic level. It is also often the case that reconfigurations are possible, in order to put the system in a functional mode that is not nominal but that still avoid a total failure (fault tolerant systems). This kind of properties could be taken into account at the operational level, with a dynamic update of the system structure and the subsequent optimal groups.

References

- [1] D.-I. Cho, M. Parlar, A survey of maintenance models for multi-unit systems, *Eur. J. Oper. Res.* 51 (1) (1991) 1–23.
- [2] R.P. Nicolai, R. Dekker, Optimal maintenance of multi-component systems: a review, in: *Complex System Maintenance Handbook* Springer, London, 2008, pp. 263–286.
- [3] L. Thomas, A survey of maintenance and replacement models for maintainability and reliability of multi-item systems, *Reliab. Eng.* 16 (4) (1986) 297–309.
- [4] S. Ozekici, Optimal periodic replacement of multicomponent reliability systems, *Oper. Res.* 36 (4) (1988) 542–552.
- [5] R.E. Wildeman, The Art of Grouping Maintenance, Ph.D. thesis, Erasmus University Rotterdam, 1996.
- [6] R. Dekker, L. Wildeman, F. Van Der Duyn Schouten, A review of multi-component maintenance models with economic dependence, *Math. Methods Oper. Res.* 45 (3) (1997) 411–435.
- [7] A. Barros, Optimisation de politiques de maintenance pour des systèmes multicomposants, in: *Mémoire de DEA Contrôle des Systèmes—Université de Technologie de Compigne*, 2000.
- [8] R. Dekker, Integrating optimisation, priority setting, planning and combining of maintenance activities, *Eur. J. Oper. Res.* 82 (1995) 225–240.
- [9] H.-C. Vu, P. Do, A. Barros, C. Berenguer, Maintenance planning and dynamic grouping for multi-component systems with positive and negative economic dependencies, *IMA J. Manag. Math.* 23 (2015) 145–170.
- [10] R.E. Wildeman, R. Dekker, A.C.J.M. Smit, A dynamic policy for grouping maintenance activities, *Eur. J. Oper. Res.* 99 (1997) 530–551.
- [11] P. Do Van, A. Barros, C. Berenguer, K. Bouvard, F. Brissaud, Dynamic grouping maintenance strategy with time limited opportunities, *Reliab. Eng. Syst. Safe.* 120 (2013) 51–59.
- [12] Y. Liu, H.Z. Huang, Optimal selective maintenance strategy for multi-state systems under imperfect maintenance, *IEEE Trans. Reliab.* 59 (2) (2010) 356–367.

- [13] M. Noureldath, D. Ait-Kadi, Optimization of series-parallel multi-state systems under maintenance policies, *Reliab. Eng. Syst. Safe.* 92 (12) (2007) 1620–1626.
- [14] R. Aggoune, Minimizing the makespan for the flow shop scheduling problem with availability constraints, *Eur. J. Oper. Res.* 115 (3) (2004) 534–543.
- [15] J. Kaabi, Y. Harrath, A survey of parallel machine scheduling under availability constraints, *Int. J. Comput. Inform. Technol.* 3 (2) (2014) 238–245.
- [16] D. Lie, Fuzzy job shop scheduling problem with availability constraints, *Comput. Ind. Eng.* 58 (4) (2010) 610–617.
- [17] B.N. Shetty, G.S. Sekhon, O.P. Chawla, K.C. Kurien, Optimization of maintenance facility for a repairable system subject to availability constraint, *Microelectron. Reliab.* 28 (6) (1988) 893–896.
- [18] M. Rausand, A. Høyland, *System Reliability Theory: Models, Statistical Methods and Applications*, second ed., John Wiley & Sons, New York, 2004.
- [19] S. Ross, *Stochastic Processes*, Wiley Series in Probability and Statistics, John Wiley and Sons, Inc., New York, 1996.
- [20] H. Pham, H. Wang, Imperfect maintenance, *Eur. J. Oper. Res.* 94 (1996) 425–438.
- [21] P. Do, A. Voisin, E. Levrat, B. Iung, A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions, *Reliab. Eng. Syst. Safe.* 133 (2015) 22–32.
- [22] J.M. Van Noortwijk, A survey of the application of Gamma processes in maintenance, *Reliab. Eng. Syst. Safe.* 94 (2009) 2–21.
- [23] I.T. Castro, A model of imperfect preventive maintenance with dependent failure modes, *Eur. J. Oper. Res.* 196 (1) (2009) 217–224.
- [24] M. Kijima, H. Morimura, Y. Suzuki, Periodical replacement problem without assuming minimal repair, *Eur. J. Oper. Res.* 37 (2) (1988) 194–203.
- [25] P.-E. Labeau, M.-C. Segovia, Effective age models for imperfect maintenance, *J. Risk Reliab.* 225 (2010) 117–130.
- [26] P. Do, H.-C. Vu, A. Barros, C. Berenguer, Maintenance grouping for multi-component systems with availability constraints and limited maintenance teams, *Reliab. Eng. Syst. Safe.* 142 (2015) 56–67.
- [27] R. Dekker, E. Smeitink, Preventive maintenance at opportunities of restricted duration, *Naval Research Logistics* 41 (3) (1994) 335–353.
- [28] E. Levrat, E. Thomas, B. Iung, Odds-based decision-making tool for opportunistic production maintenance synchronization, *Int. J. Prod. Res.* 46 (2008) 5263–5287.
- [29] L. Radouane, C. Alaa, A. Djamil, Opportunistic policy for optimal preventive maintenance of a multi-component system in continuous operating units, *Comput. Chem. Eng.* 9 (9) (2009) 1499–1510.
- [30] A. Van Horenbeek, L. Pintelon, A dynamic predictive maintenance policy for complex multi-component systems, *Reliab. Eng. Syst. Safe.* 120 (2013) 39–50.
- [31] A. Bouvard, S. Artus, C. Berenguer, V. Cocquempot, Condition-based dynamic maintenance operations planning and grouping. Application to commercial heavy vehicles, *Reliab. Eng. Syst. Safe.* 96 (6) (2011) 601–610.
- [32] P. Do Van, A. Barros, C. Berenguer, Reliability importance analysis of Markovian systems at steady state using perturbation analysis, *Reliab. Eng. Syst. Safe.* 93 (11) (2008) 1605–1615.
- [33] H.-C. Vu, P. Do, Barros, A stationary grouping maintenance strategy using mean residual life and the Birnbaum importance measure for complex structures, *IEEE Trans. Reliab.* 65 (1) (2016) 217–234.

Reliability estimation considering warranty data, usage rate profile, and qualitative information on design changes

9

S. Limon, O.P. Yadav

North Dakota State University, Fargo, ND, United States

1 Introduction

Reliability is the measure of the time-dependent quality of a product and is considered as a critical factor to the success of any business and customer satisfaction. The reliability measures the product's performance, capability, and durability over a period of time considering uncertainty caused by several factors, such as variation in design, material, manufacturing, and operating conditions. According to O'Connor [1], reliability is "the probability that an item will perform a required function without failure under stated conditions for a stated period of time." As an assurance of the product reliability, manufacturers provide warranty packages to the customer during a sale, which is designed keeping in mind the product reliability estimates. The manufacturers provide different warranty packages based on product types, geographic areas, and several other business factors. Regardless of the packages and policies, the warranty is always a liability that incurs a cost to the manufacturers by means of money and reputation, and for customers, it causes sometimes a big financial loss with downtime and casualty. For example, the US manufacturers alone spend more than 25 billion US dollars per year to resolve the warranty-related issues [2]. It is, therefore, important to provide a more realistic and accurate reliability estimate of the product to minimize these huge costs and the casualties.

In reliability analysis, the field failure data provide more authentic information about the actual product behavior compared to laboratory data [3,4], because field data capture the product usage behavior combined with the environmental exposures a product has gone through during its actual use [5]. The failure data during the warranty period mostly known as the field warranty data represent a great source and readily available modes of field data. However, there are several drawbacks with this warranty data, such as the sparse nature of data, incomplete, unclear, and delays and mistakes in reporting [5]. Due to these drawbacks, the reliability estimates based on only failure data will usually be biased and inaccurate. Considering these issues, it is important to develop a more effective reliability assessment method that captures maximum available information related to both failed and unfailed population that will provide more accurate reliability estimation.

Several efforts have been made in past to capture information on unfailed population. The supplementary follow-up data collection guidelines have been provided by Kalbfleisch and Lawless [6]. The follow-up study is used to collect product usage information of a fraction of the unfailed population and a pseudo-likelihood estimation approach is developed considering these follow-up data. Both parametric and nonparametric methods were attempted to estimate reliability [6–9]. The non-homogeneous Poisson process is considered by Suzuki [10] whereas Alam and Suzuki [11] used only failure data assuming the censored usage time is unknown. Oh and Bai [4] considered field failure data collected even after the warranty period to estimate field failure and lifetime distribution. Nonetheless, gathering this information through follow-up study is costly, time-consuming, and provides partial information regarding the unfailed population. Furthermore, sometimes product users are not willing to share all relevant information that make the lifetime and reliability estimation considering unfailed population a very difficult exercise.

This chapter builds on the work done by Limon et al. [12] and Yadav et al. [13] to provide a comprehensive framework that utilizes warranty/failure data, censored data for surviving population derived from the usage rate-based approach [12] and considering impact of new design changes in product reliability [13]. For capturing and estimating usage rate information, different kinds of field data, such as warranty claims data, maintenance data, and other follow-up data, for example, recall data, survey data, and other supplementary data if available, have been utilized. The censored data (accumulated usage) are generated for unfailed population considering the usage rate and age of the product. For capturing the impact of design changes on product reliability, the fuzzy logic approach is used to quantify the qualitative information using appropriate reliability improvement indices as suggested in Yadav et al. [13]. Finally, a reliability estimation approach is demonstrated for a real-time case example assuming lifetime follows the Weibull distribution.

2 Estimation of censored data for surviving population

2.1 Usage rate profile

The measure of usage or usage rate mostly depends on product types. For example, the mileage of an automobile, the number of copies of a photocopy machine, and the operation hours of utility equipment might be considered as usage for each respective product types. During the usage rate estimation, one needs to gather product usage-related information from the field or market data. However, the collection of usage-related information for the entire population is almost impossible. Additionally, in existing literature, it is assumed that the accumulated usage for any given product is different for failed and censored population, and in most of the cases the accumulated usage of the censored population is unknown [11]. Though, in few cases the severe failures might affect the usage rate to some extent; in general, the usage rate is independent of failure modes but it mostly depends on the customers' usage

behavior. We therefore assume that the usage rate will be the same for failed and surviving population. Furthermore, Lawless et al. [14] also made an assumption that mileage accumulation rate is independent of failure time that supports our assumption too.

The product usage data can be collected from several sources, such as recall data, maintenance data, warranty data, and online connect data. Partial surveys and recall data have been used as a part of censored data to estimate the parameters of field performance [6,8]. Regular maintenance data can be collected from dealers or maintenance department where customers bring their product for regular maintenance during the product life cycle that includes after the warranty period. It is important to keep in mind that this information is for the unfailed product because it is collected during regular maintenance work. Another source of usage data is warranty claims database. Since warranty claims represent only failure data during the warranty period, it provides relatively better customer usage information, the number of hours accumulated, and other types of usage data that may be utilized to estimate usage rate of the product. Though warranty data have several shortcomings, they contain a great source of information regarding the actual product performance. Another source of customer usage data is the online connecting data. As technology grows, it is possible to track the usage of a product utilizing a microchip to capture the real-time utilization. This approach is expensive and sometimes the customers might not allow tracking their usage behavior, but it provides possible mode of collecting usage information in many cases [15,16]. Production and sales data can be used to capture the time that a product is in service, also known as product age. In order to estimate the usage rate, it is essential to gather both the accumulated usage and the time in service data, accurately.

In order to process the data, collected from several sources, it is important to identify the variables of interest, such as product age, accumulated usage, product model, and other related variables. The outliers (extremely large or small and infeasible data points, for instance, negative product age) in each category of variables should be removed. Also, it is necessary to screen out all other data points or variables that do not match with the product model under consideration. After gathering all relevant data, the usage rate is calculated considering the accumulated usage from the available data and product age (or time in service). Once all the variables of interests are in hand, the usage rate is calculated. The usage rate u_i for the i th product out of n number of products is obtained as,

$$u_i = \frac{w_i}{t_i}, \quad i = 1, 2, \dots, n \quad (1)$$

where w_i and t_i represent total usage and time in service for the i th product, respectively. It is important to note that both the total usage and the service in time will always be positive; therefore, the usage rate also must be a positive value varying from zero to any real value. It is important to note that the total usage for any given product can be obtained from warranty claims, maintenance and service records, and any other

sources available to the manufacturers. The time in service information can easily be obtained from the sales records of the dealers. The next section provides discussion on the usage rate distribution.

2.2 Usage rate distribution

In this work, a parametric model is assumed for the usage rate distribution. Since the usage rate varies from customer to customer, it is important to treat it as a random variable and establish an appropriate usage rate distribution. Earlier studies [9,17] show that for automobiles the usage rate is generally linear over time and follows the lognormal distribution. The usage rate data histogram of an industrial utility equipment shows the lognormal distribution as a better fit with the lower Anderson-Darling (AD) statistics (see Fig. 1), which essentially supports the earlier assumption on usage rate distribution. Furthermore, the usage rate also varies between different market segments with different geographical areas that also supports our argument of treating the usage rate as a random variable.

Therefore, the lognormal distribution is a strong candidate to model the random behavior of the usage rate. Considering a random variable U that follows a distribution with probability distribution function (pdf) $f(u)$, the likelihood function will be given as

$$L(\theta) = \prod_{i=1}^n f(u_i, \theta) \quad (2)$$

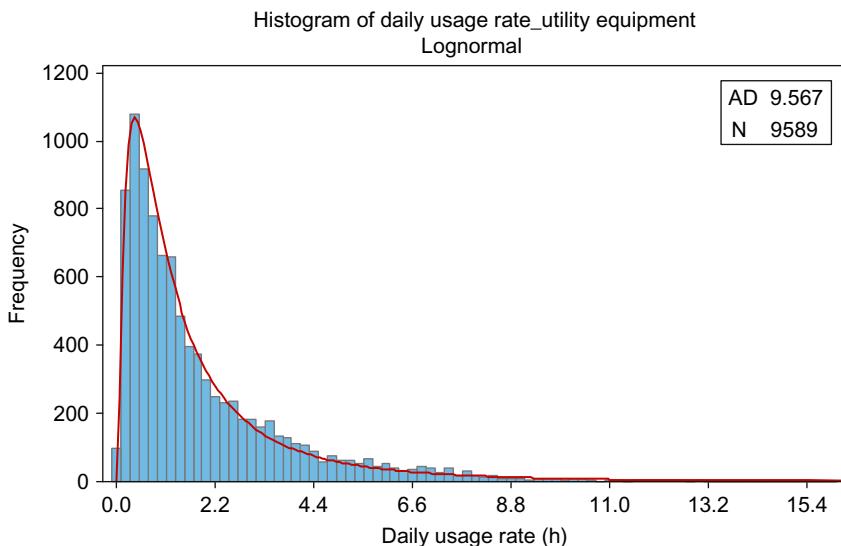


Fig. 1 Histogram of usage rate with the lognormal fit of utility equipment.

By taking the natural logarithm on both sides of Eq. (2), the log-likelihood function is written as

$$\log L(\theta) = \sum_{i=1}^n \log f(u_i, \theta) \quad (3)$$

where θ represents the parameter of interests that need to be estimated. For the lognormal distribution, the parameters are known as the location and the scale $\theta=(\mu, \sigma)$, respectively. The maximum likelihood estimates of these two parameters were given as

$$\hat{\mu}_u = \frac{\sum_{i=1}^n \log(u_i)}{n} \quad (4)$$

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n (\log(u_i) - \hat{\mu}_u)^2}{n} \quad (5)$$

Generally, the distribution model is fitted based on the available data to estimate model parameters. However, the usage rate data gathered usually represents a small fraction of the surviving population and the estimation of model parameters based on this small fraction of the population will have higher uncertainty. It is, therefore, important to use a parametric bootstrap method [18,19] for estimation of model parameters. The bootstrap provides a robust estimation of model parameters with resampling method. The bootstrap resampling algorithm for estimating the usage rate parameters works as discussed. First, an available fraction of population usage rate will be chosen as a bootstrap initial sample, $x=x_1, \dots, x_n$. Then, start resampling from the initial sample x with replacement and index each of the sample replicates by $b=1, \dots, B$, where B is a very large number. Each generated sample size will remain same as the initial sample size n . After that, compute the estimate of the interested parameters $\hat{\theta}(b)$ for each bootstrap sample generated during earlier steps. Finally, take the sample average of all sample parameter estimates $\hat{\theta}$ as given below:

$$\hat{\theta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}(b) \quad (6)$$

To get the confidence interval of bootstrap estimates, it is necessary to estimate the standard error of the bootstrap resampling estimates using equation given below:

$$\hat{s}\hat{e}(\hat{\theta}^*) = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}(b) - \hat{\theta}^*)^2} \quad (7)$$

Final bootstrap estimates of model parameters are then used to generate censored usage data for the surviving population. The next section provides a detailed discussion on generating the accumulated usage time for surviving populations.

2.3 Estimation of accumulated usage of censored data

Once the usage rate distribution parameters were estimated, the next step is to generate censored data for the surviving population using these parameter estimates. Assuming at any given point of time, the age of the product (or time in service) is also defined as a random variable; then the accumulated total usage is given as,

$$T_c = U \times A \quad (8)$$

where T_c denotes the censored usage time; U and A represent random variables of the usage rate and the product age, respectively. The product age can be found from the product manufacturing and sales-related information, which is readily available in the warranty database or with dealers. The measurement unit of A is calendar time such as days, weeks, or months, whereas usage rate is measured as usage per calendar time such as mileage per day or usage hours per day.

Since the accumulated usage (or censored) time T_c is a product of two random variables, it is also treated as a random variable. Considering usage rate and product age (time in service) as two independent variables, the expected value and variance of the accumulated usage time T_c can be determined by the following equations [20]:

$$E(T_c) = E(U) \times E(A) = \mu_U \mu_A \quad (9)$$

$$\text{Var}(T_c) = \text{Var}(U)\text{Var}(A) + \text{Var}(U)(\mu_A)^2 + \text{Var}(A)(\mu_U)^2 \quad (10)$$

If the distribution types, model parameters of usage rate, and product age were known, it becomes easier to generate censored data for the surviving population. There are two ways to get the censored population age. First one is getting the actual age of each product from manufacturing sales database and the second way is to treat the product age as a random variable. Considering the actual age of individual units A_i , the accumulated usage time is calculated as,

$$T_{c,i} = U_i \times A_i \quad (11)$$

where the random variable usage rate is assumed to follow the lognormal distribution $U_i \sim \log n(\hat{\mu}_u, \hat{\sigma}_u^2)$ and A_i is the actual age of the i th unit.

In the second case, we treat the population age as a random variable that follows a specific probability distribution function. The motivation for considering product age as a random variable is that in many real-life cases the actual age of each item is not known or sometimes very difficult to obtain. Generally, units are produced based on product demand or production capacity and subsequently end up in the field after selling them to customers. For example, auto companies produce a certain number of

units every month and sell those units to customers in the market. At the same time, a certain amount of units would have spent enough time in the market and will be getting out of the warranty period. If we visualize this continuous process of a certain number of units being produced and getting into the market per unit time (week or month) and also almost a similar number of units were going out of the warranty period, it almost represents a steady flow process where on an average there are a similar number of units moving through the system. We believe this scenario can be reasonably modeled as a uniform distribution function where one parameter represents warranty time length and the other parameter captures the average number of units entering and/or leaving the warranty period.

Considering that the distribution of usage rate follows the lognormal distribution $U_i \sim \log n(\hat{\mu}_u, \hat{\sigma}_u^2)$ and the product age follows the uniform distribution function $A_i \sim \text{unif}(b, c)$, Eq. (11) can be used to generate the censored time data. The parameters for the uniform distribution are given as

$$E(A) = \frac{b+c}{2} \quad (12)$$

$$\text{Var}(A) = \frac{(c-b)^2}{12} \quad (13)$$

Using the distribution parameters of these two distributions, random datasets can be generated for both random variables. The number of data points in each random dataset should be equal to the surviving population size. These two random number datasets can then be used to generate censored time for the surviving population using Eq. (11). Once censored data are available, the maximum likelihood method can be applied to estimate model parameters of the combined dataset. Alternatively, given that the distribution parameters for both random variables are known, Eqs. (9), (10) can be used to estimate the parameters of the accumulated usage (censored) time. These estimated parameters can then be used to generate censored data for the surviving population. To understand the actually accumulated usage data and the overall reliability, the next section provides the estimation guidelines in detail.

3 Field reliability estimation

3.1 Estimation model

The Weibull distribution is assumed to represent the lifetime of a product as it is widely suggested in the literature [19]. The probability distribution and the survival (reliability) function of a Weibull random variable are given as follows:

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta} \right)^{\beta-1} e^{-(t/\eta)^\beta} \quad (14)$$

$$S(t) = \bar{F}(t) = 1 - F(t) = e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (15)$$

where β and η are the shape and scale parameters, respectively. To assess the product reliability, estimation of distribution parameters is the most critical step in a warranty data analysis, especially when the usage data were acquired from different sources and surviving data were generated from estimation.

The simplest scenario arises when there are only failure data and no censored data. The maximum likelihood function for only failure data is given as

$$L(\theta) = \prod_{i=1}^r f(t_i, \theta) \quad (16)$$

Here r is the number of failures. However, the necessity to provide a more realistic reliability estimates requires not only failure data but also to capture the maximum available information related to the surviving population. The inclusion of this information as censored data results in a more complex data analysis problem. Considering the Weibull life distribution, the likelihood function is given as,

$$L(\theta) = \prod_{i=1}^r f(t_i, \theta) \prod_{j=r+1}^n S(t_{cj}, \theta) \quad (17)$$

where r is the number of failures, n is the total number of data points, θ is the parameter of interest, and $f(t_i, \theta)$ and $S(t_{cj}, \theta)$ are the probability distribution function and the survival function, respectively. Taking logarithm in both sides of Eq. (17) after using the probability density function and survival function of the Weibull distribution given in Eqs. (14)–(15), the log-likelihood function is given as

$$\log L(\theta) = r \log \beta - r \log \eta + (\beta - 1) \sum_{i=1}^r \log \left(\frac{t_i}{\eta} \right) - \sum_{i=1}^r \left(\frac{t_i}{\eta} \right)^\beta - \sum_{j=r+1}^n \left(\frac{t_{cj}}{\eta} \right)^\beta \quad (18)$$

To find the estimates of model parameters, β and η , we need to maximize the log-likelihood function of Eq. (18). However, it is impossible to achieve closed-form solution of Eq. (18), and therefore, necessary to solve the log-likelihood function by using an appropriate numerical method. Modern statistical software *R* is used to find the MLEs (maximum likelihood estimates) by numerical method. A nonlinear built-in optimization function *optim* that is based on an algorithm provided by Nelder and Mead [21] is used to maximize the log-likelihood function.

Once the distribution parameters were estimated, the reliability can easily be estimated by using Eq. (15). It is important to note that censored usage time information is to be used to update the distribution parameters for providing a more realistic field reliability estimation.

3.2 A case example of field reliability estimation

The real-time case example is presented to demonstrate the methodology of generating the censored data for surviving population and estimating the product reliability. We consider the field data of utility equipment used in construction, maintenance, agriculture, and other application areas. The manufacturer provides a 12-month warranty with unlimited usage hours. The product is launched into the market starting in early 2009 and all the claims up to the end of 2013 were recorded. It is assumed that all failures within the warranty period have been reported and nonreported items were considered as censored population. The available field data include warranty claims, maintenance, and recall data. Each type of field data contains various usage-related information, such as the date of sale, date of failure (maintenance or recall), and accumulated machine hours. Many of the claims were out of the warranty period though they were also considered in estimating the usage rate to increase the sample size and get better estimates as well.

A data screening process was conducted for removing all infeasible or outlier data, such as negative product age, same data point recorded for multiple failure modes, daily usage rate more than 24 h, and extremely low usage rate. The final dataset contains a total of 9004 field claims with 6570 of the claims within the warranty period, 1120 claims beyond the warranty period, and 1314 recall and maintenance data. These refined field data were then used to calculate the usage rate for each individual unit using Eq. (1) considering the accumulated usage (machine) hours and age of the product. The collected usage rate data provided a good fit to the lognormal distribution as specified by a smaller value of the AD (8.289) metric (see Fig. 2), which supports our

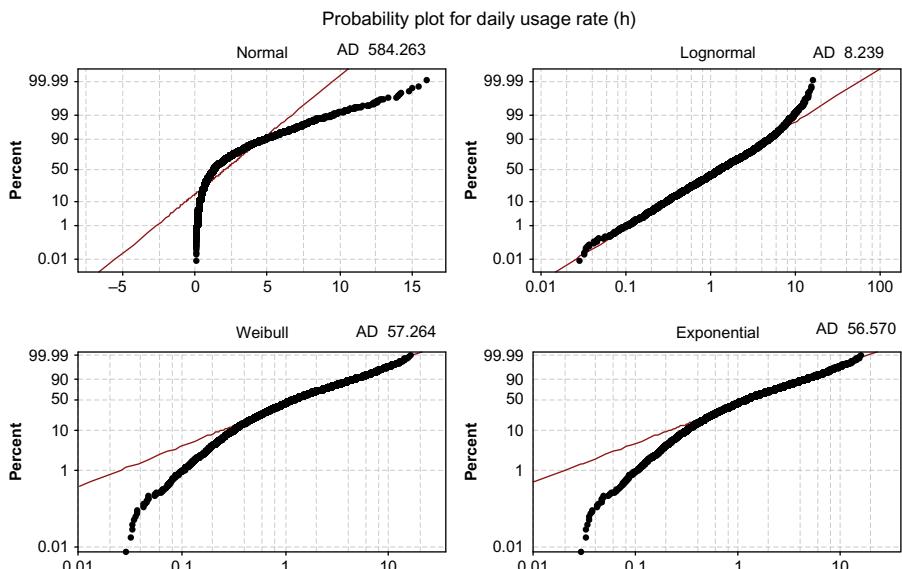


Fig. 2 Probability plot for usage rate distribution.

initial assumption regarding the usage rate distribution. The individual unit usage rate of this fraction of the population is then used to estimate the usage rate distribution parameters representing the entire population. The bootstrap resampling method is applied to get more robust parameters of the usage rate distribution. [Table 1](#) shows the estimated parameters of the usage rate distribution based on original sample and bootstrap resampling. These estimated parameters are then used to generate censored usage time data for the surviving population. For generating censored usage time, we have considered the units manufactured and sold during the year 2011 only. During the year 2011, there were 780 claims recorded in the warranty data out of the total 2636 individual units sold and operational in the field. Using usage rate distribution parameters estimated earlier (given in [Table 1](#)), the censored usage time is generated for the remaining 1856 unfailed units.

Considering both the population age and usage rate as two random variables following different distributions, the censored usage time data were generated. Once the censored dataset for the unfailed population is available, lifetime parameters were estimated considering both failure and censored data assuming the lifetime follows the Weibull distribution. Realizing that the MLE equation ([Eq. 18](#)) does not provide closed-form solutions, the statistical software *R* is used that applies a numerical method to solve the MLE. [Table 2](#) shows the estimated Weibull parameters. Using these parameters and the survival function equation, system reliability is estimated

Table 1 Estimated usage rate parameter values

Usage rate distribution	Parameters	Estimate	Lower 95% CI	Upper 95% CI
Lognormal (bootstrap)	$\hat{\mu}_U$	0.1821	0.1607	0.2034
	$\hat{\sigma}_U$	1.0361	1.0221	1.0500
Lognormal (original)	$\hat{\mu}_U$	0.1822	0.1608	0.2035
	$\hat{\sigma}_U$	1.0360	1.0209	1.0510

Table 2 Estimated lifetime parameters considering the Weibull distribution

Weibull analysis	Parameters	Estimate	Lower 95% CI	Upper 95% CI
Only failure	$\hat{\beta}$	0.6645	0.6289	0.7000
	$\hat{\eta}$	84.39	74.97	93.81
Actual age	$\hat{\beta}$	0.4874	0.4574	0.5163
	$\hat{\eta}$	2229.99	1810.95	2649.03
Distribution age	$\hat{\beta}$	0.4817	0.4534	0.5101
	$\hat{\eta}$	2248.61	1827.26	2669.97

for different usage times. Fig. 3 shows the reliability behavior of the system for different scenarios discussed in the chapter.

Table 2 shows the Weibull parameters for both failure data and the combined dataset indicating early failure issues. The estimated results show that for all cases the Weibull shape parameter value is approximately 0.50, which implies an infant mortality rate ($\beta < 1$). The immature product design (PD) might be the reason for this infant mortality. As this analysis was carried out within the second year of the product launch, there is very high possibility that the design still has some deficiencies. As given in Table 2, it shows that the characteristic life has been significantly improved after the inclusion of the surviving population information into the analysis. This indicates the advantage of incorporating surviving population-related information for obtaining more realistic reliability estimates. However, the inclusion of additional information into the reliability analysis does not change the shape parameter of the Weibull distribution, which conforms to the shape parameter property. Table 3 shows the results of different simulation studies. The number of failures predicted based on the proposed approach is very close to the actual number of failure. This reinforces our argument and assumption that inclusion of unfailed population-related information in reliability analysis provides more realistic estimates.

The case example results clearly show that the reliability assessment based on only failure data is unrealistic and biased. Fig. 3 shows a big gap between the reliability estimates based on only failure data and with the inclusion of the surviving population data. Our further investigation on the process of generating surviving data highlights the impact of random variable product age (A) on the reliability assessment. As shown in Fig. 3, the distribution age line, wherein product age variable follows the uniform distribution, provides a very close reliability estimates to actual age reliability line wherein we considered the actual age of each individual unit in the field. This certainly

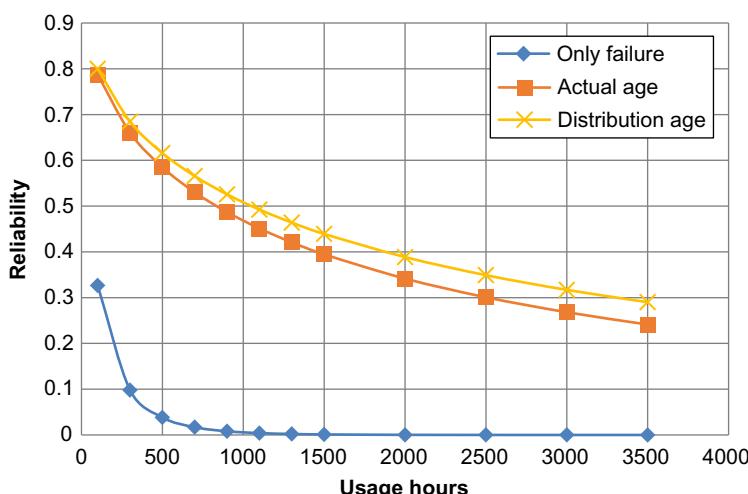


Fig. 3 Reliability comparison among different approaches.

Table 3 Simulation results with actual and predicted failures

Total population	20% Failure			30% Failure			40% Failure		
	Actual	Predicted	Error (%)	Actual	Predicted	Error (%)	Actual	Predicted	Error (%)
2500	500	489	2.20	750	759	1.20	1000	1038	3.80
10,000	2000	1958	2.10	3000	3038	1.26	4000	4157	3.92
20,000	4000	3927	1.82	6000	6076	1.26	8000	8325	4.06
100,000	20,000	19,624	1.88	30,000	30,498	1.66	40,000	41,785	4.46

cautions practitioners not to spend critical resources in extracting actual age data from the database but motivates them to consider a suitable distribution of population age to obtain a realistic reliability assessment.

4 Impact of design change on reliability improvement

4.1 Fuzzy logic in reliability

Previous sections presented the usage-based field reliability estimation considering both the failure and surviving data, which provides the estimates for the existing PD. As a part of innovation and continuous improvement, the manufacturers make changes into their existing PD, material, and manufacturing processes. During the product development and design stages, it is necessary to evaluate the reliability of a new or improved design to assess the improvement from their previous designs. Since the new design is still under the development process, there exist no field failure or test data to assess the product reliability using the traditional reliability assessment methods.

To overcome these problems, a fuzzy logic-based reliability estimation is adopted during the PD process [13]. In early stages of the PD process, there exist enough qualitative information from different sources, such as physics-of-failure models, results from robust design experiments, computer-aided engineering (CAE) analysis, failure modes and effect analysis, and pure expert opinions among many others. The qualitative information is captured as an engineering judgment or an improvement index, which represents an improvement of a particular failure mode or mechanism. These improvement indices were treated as fuzzy numbers or linguistic variables, combined and processed through a fuzzy reasoning method to map their impact on product reliability. The crisp output of the fuzzy reasoning method is treated as a new evidence. Then, the Bayesian approach is used to integrate the new evidence with other types of information such as quantitative data and qualitative information of a PD.

4.2 Input-output variables identification

The input variables are improvement indices representing the impact of design changes and corrective actions on identified failure modes. The engineering judgments are defined as improvement indices (I_{XX}) as follows,

$$I_{XX} = \frac{(\text{Gain})}{(\text{Previous})} \times 100 \quad (19)$$

where the subscript XX indicates the specific failure mode or mechanism, (Gain) represents a net improvement in failure mode or functional characteristic, and (Previous) denotes the level of functional characteristic before the corrective actions or design changes were made. For instance, past failure analysis identifies three major failure modes of a hydraulic system: noise, fluid leakage, and fatigue failure. The improvement indices for these three failure modes were described below in detail.

Noise: Sometimes design of experiments or robust design experiments are used to optimize a design or address the identified failure issues. Depending on the failure mode (noise in this case), an appropriate measurable characteristic is identified and optimized using designed experiments. The gain or improvement in that measurable characteristic is used to capture the improvement index. If a robust design experiment is used, which uses the signal-to-noise (SN) ratio as a measurable indicator to find the optimal solution, the improvement in a noise problem can be captured as follows:

$$I_{NOISE} = \frac{(SN_a - SN_b)}{SN_b} \times 100 \quad (20)$$

where subscripts, a and b , represent the signal-to-noise ratios after and before corrective actions were taken based on the experiment outputs.

Leakage: Where there are no measurable characteristics available to measure an improvement due to design changes, the information from the failure mode and effect analysis (FMEA) documents can be used. When design team captures and documents the changes made in the design, it is required to estimate the impact of design changes by revising the risk priority number (RPN) that defines the severity and occurrence propensity of each failure modes. Assuming that there are no other sources to capture the improvement on leakage problem, the revised RPN values are the only source of quantitative data that can measure the leakage conditions. The improvement of a fluid leakage problem can be estimated by comparing the RPN values of two different designs as follows:

$$I_{LEAK} = \frac{(RPN_b - RPN_a)}{RPN_b} \times 100 \quad (21)$$

Similar to noise problem, subscripts a and b represent the RPN values after and before the corrective actions were taken based on the FMEA documents.

Fatigue failure: The improvement of a fatigue failure problem can be measured by comparing the fatigue life. The CAE analysis and use of an appropriate physics-of-failure model can provide the fatigue life improvement of a product or component, which can be captured as follows:

$$I_{FF} = \frac{(N_a - N_b)}{N_b} \times 100 \quad (22)$$

where subscripts a and b represent the fatigue life after and before the design changes.

The change in the system reliability parameter due to corrective actions was considered as an output variable and termed as the reliability improvement index (RII). The RII can be estimated by Eq. (23)

$$RII = \frac{(\lambda_b - \lambda_a)}{\lambda_b} \times 100 \quad (23)$$

Here λ represents the transformed failure rate. The subscripts a and b indicate the failure rate after and before any changes were made.

4.3 Fuzzy linguistic variables for input-output variables

The triangular membership function is assumed to estimate the improvement of any failure category in linguistic terms. These fuzzy categories are defined as negligible, very low, low, moderate, high, and very high. The fuzzy sets or the term sets $T_{I_{XX}}$ for each input (improvement index I_{XX}) are defined by Eq. (24) whereas the fuzzy sets or term sets T_{RII} for output (RII) are defined by Eq. (25). The membership function is graphically shown in Fig. 4.

$$T_{I_{XX}} = \{\text{negligible, very low, low, moderate, high, very high}\} \quad (24)$$

$$T_{RII} = \{\text{negligible, very low, low, moderate, high, very high}\} \quad (25)$$

4.4 Fuzzy rules, reasoning, and defuzzification

The failure information and the expert opinion can be utilized to develop the fuzzy IF-THEN rules. The field warranty data analysis along with expert opinion could be of great help in building the fuzzy IF-THEN rules because this analysis represents the actual system behavior in the field. The example of fuzzy IF-THEN rules is given below:

IF I_{NOISE} is negligible and I_{LEAK} is negligible and I_{FF} if negligible THEN RII is negligible
IF I_{NOISE} is negligible and I_{LEAK} is very low and I_{FF} if moderate THEN RII is low

These fuzzy IF-THEN rules build a fuzzy system that converts the fuzzy inputs into a fuzzy output. The fuzzification of the input is the initial step of the fuzzy reasoning processes because usually the inputs are available in a crisp value. After the inputs are fuzzified, it is necessary to determine the satisfying fuzzy rules. The given values of inputs fire the IF-part of all rules in parallel and to some degrees.

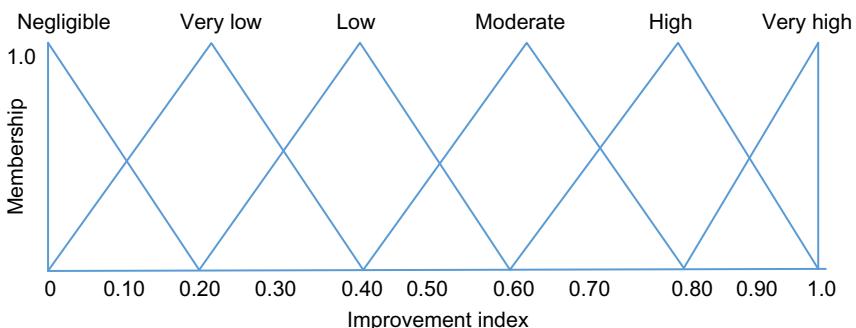


Fig. 4 Membership function for improvement indices and output variables.

The AND or min T-norm operator is used when there is more than one input. The matching (combination) of different values of linguistic variables in the premise is achieved through the fuzzy IF-THEN rules as follows:

$$\mu_{B'} = \min \left\{ \mu_{A'_1}(x), \mu_{A'_2}(x), \dots, \mu_{A'_n}(x) \right\} \quad (26)$$

The fuzzy sets that represent the output of each active/live rule are then aggregated into a single fuzzy set with the help of the max S-norm operator using the following equation:

$$\mu_B(y) = \max \{ \mu_{B^1}(y), \mu_{B^2}(y), \dots, \mu_{B^m}(y) \} \quad (27)$$

Since the output of the fuzzy reasoning process is also a fuzzy set, it needs to be defuzzified and converted into a crisp value. The centroid of area (Z_{COA}) method can be applied for the defuzzification process using Eq. (28)

$$Z_{COA} = \frac{\int \mu_B(Z) Z dZ}{\int \mu_B(Z) dZ} \quad (28)$$

where $\mu_B(Z)$ is the aggregated output membership function.

4.5 Fuzzy reliability analysis

Assuming the life follows the Weibull distribution, a transformed failure rate (λ) is defined by Eq. (29) as given below:

$$\lambda^* = \frac{1}{\eta^\beta} = \frac{1}{T^*} \quad (29)$$

After defuzzification, the output of the fuzzy inference model provides a crisp value of RII. This RII value is then converted into a transformed failure rate ($\lambda_t = 1/T^*$) using Eq. (23). This point estimate of the transformed failure rate (λ_t) treated as a new information or evidence and offered to the Bayesian framework [22] to update reliability parameters after the corrective actions were taken as follows:

$$f\left(\frac{\lambda_t}{T^*}\right) = \frac{L\left(\frac{T^*}{\lambda_t}\right) f(\lambda_t)}{\int L\left(\frac{T^*}{\lambda_t}\right) f(\lambda_t) d\lambda_t} \quad (30)$$

where $f(\lambda_t/T^*)$ is the posterior probability density function of the transformed failure rate (λ), $L(T^*/\lambda_t)$ is the likelihood function, $f(\lambda_t)$ is the prior probability density

function of the transformed failure rate, T^* is referred as the new evidence of the transformed failure rate or mean time to failure.

4.5 A case example of updating the reliability using fuzzy logic

In earlier sections, the Weibull parameters were estimated considering both the failure and surviving population. The Weibull parameter values were rounded off for simplicity of the estimation (shape = 0.50, characteristic life = 2250). It is assumed that the shape parameter of the Weibull distribution will remain same and is known here (through estimation). In this case, the transformed failure time t^β follows an exponential distribution [22]. The assumption of known shape factor and consideration of the exponential distribution allow selecting the gamma distribution as the prior probability distribution for the Bayesian analysis.

The Weibull parameters were estimated earlier based on the field failure data and surviving population. Then the initially transformed failure rate (λ_0) is estimated using the Weibull parameters as follows:

$$\lambda_0 = \frac{1}{\eta^\beta} = 2.10 \times 10^{-2}$$

The mean time to failure (MTTF) can be estimated by inverting the failure rate (λ). This failure rate is considered as the failure rate of the old design. According to our assumption,

$$\lambda_0 = \frac{a_0}{b_0} = 2.10 \times 10^{-2}$$

or,

$$a_0 = b_0 (2.10 \times 10^{-2})$$

Here, a and b represent the gamma distribution parameters. The prior parameters of the gamma distribution can be estimated by solving the following relation:

$$\int_{0.2\lambda_0}^{8\lambda_0} b_0^{a_0} \lambda^{a_0-1} e^{-b_0\lambda} d\lambda = 0.99$$

The above integral is not in a simple form; thus, the numerical integration has been applied (Matlab or R software) that provides the prior parameters as $a_0 = 3.90$ and $b_0 = 1.86 \times 10^2$. **Table 4** shows the prior parameters and reliability information of the old design.

Now, we assume that earlier failure issues have been addressed and the new design is an improved version of the earlier (old) design. Using improvement indices identified earlier for noise, leakage, and fatigue failure, the estimated improvement indices were given as $I_{\text{NOISE}} = 0.65$, $I_{\text{LEAK}} = 0.50$, and $I_{\text{FF}} = 0.35$, respectively. The output of

Table 4 Reliability information of the old design

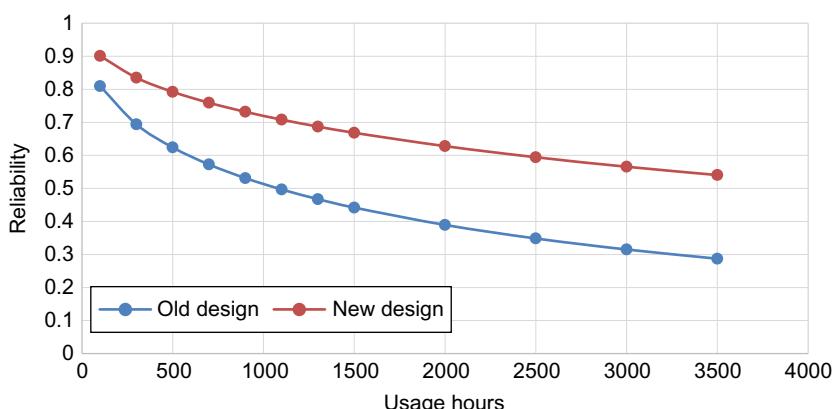
Transformed failure rate (λ_0)	Transformed mean time to failure (MTTF)	Prior gamma distribution	
		a_0	b_0
2.10×10^{-2}	47.61	3.90	1.86×10^2

fuzzy reasoning process, based on the centroid of area method (Eq. 26), provides the crisp value of RII as 0.501. The RII value is then converted into transformed failure rate of the new design using the Eq. (21) and treated as a new information or evidence. This new information is further used to get the posterior gamma distribution parameters of the Bayesian analysis. The characteristic life (η) of the new design has been also estimated according to the Bayesian analysis results.

Table 5 provides the posterior gamma distribution parameters and transformed failure rate of the new and improved design. It is found that the transformed failure rate has been improved (reduced) in the new design after taking corrective measures. This also implies the improvement of the characteristic life of the Weibull distribution parameter that can be estimated by using Eq. (29). Fig. 5 shows the improvement

Table 5 Reliability parameters of the new design

RII	Transformed failure rate (λ_0)	Transformed mean time to failure (MTTF)	Posterior gamma distribution	
			a_0	b_0
0.501	1.04×10^{-2}	95.42	3.90	3.75×10^2

**Fig. 5** Reliability comparison of the old and new design.

and comparison in the reliability of old and new design at different life cycle after the corrective actions were taken. For further understanding on fuzzy reliability assessment, authors are advised to refer Yadav et al. [13].

5 Conclusion

This work presents a reliability estimation framework considering both the surviving population and failure data by utilizing the customer usage profile. It further captures the impact of the design changes in reliability estimation by using the fuzzy logic approach. The case study results show that the inclusion of the surviving population significantly improves the accuracy of the reliability estimation process. The framework also allows to further update the reliability estimates if any changes were made in the design to address the failure issues or new design is suggested based on the previous model. The comprehensive reliability assessment framework allows to utilize all sorts of information available or generated during the product development process. The major challenge for this framework is to gather more accurate information related to the surviving population and to generate censored time using this information. However, the advanced communication and data management technology provides a great help to gather more accurate information. The incorporation of the fuzzy logic to aggregate the qualitative information and the Bayesian analysis method ensures the robustness of the reliability estimates of the proposed approach.

References

- [1] P.D.T. O'Connor, Practical Reliability Engineering, fourth ed., John Wiley & Sons, Ltd, Chichester, England, 2002.
- [2] D.K. Manna, S. Pal, S. Sinha, A use-rate based failure model for two-dimensional warranty, *Comput. Ind. Eng.* 52 (2007) 229–240.
- [3] M.R. Karim, K. Suzuki, Analysis of warranty data with covariates, *Proc. IMechE Part O: J. Risk and Reliab.* 221 (2007) 249–255.
- [4] Y.S. Oh, D.S. Bai, Field data analysis with additional after-warranty failure data, *Reliab. Eng. Syst. Saf.* 72 (2001) 1–8.
- [5] B. Rai, N. Singh, Hazard rate estimation from incomplete and unclean warranty data, *Reliab. Eng. Syst. Saf.* 81 (2003) 79–92.
- [6] J.D. Kalbfleisch, J.F. Lawless, Estimation of reliability in field performance studies, *Technometrics* 30 (4) (1988) 365–378.
- [7] K. Suzuki, Nonparametric estimation of lifetime distributions from a record of failures and follow-ups, *J. Am. Stat. Assoc.* 80 (389) (1985) 68–72.
- [8] K. Suzuki, Estimation of lifetime parameters from incomplete field data, *Technometrics* 27 (3) (1985) 263–271.
- [9] X.J. Hu, J.F. Lawless, K. Suzuki, Nonparametric estimation of a lifetime distribution when censoring times are missing, *Technometrics* 40 (1) (1998) 3–13.
- [10] K. Suzuki, Analysis of field failure data from a nonhomogeneous Poisson process, *Rep. Stat. Appl. Res.* 34 (1987) 6–15.

-
- [11] M.M. Alam, K. Suzuki, Lifetime estimation using only failure information from warranty database, *IEEE Trans. Reliab.* 58 (4) (2009) 573–582.
 - [12] S. Limon, O.P. Yadav, M.J. Zuo, J. Muscha, R. Honeyman, Reliability estimation considering usage rate profile and warranty claims, *Proc. IMechE Part O: J. Risk and Reliab.* 230 (3) (2016) 297–308.
 - [13] O.P. Yadav, N. Singh, R.B. Chinnam, P.S. Goel, A fuzzy logic based approach to reliability improvement estimation during product development, *Reliab. Eng. Syst. Saf.* 80 (2003) 63–74.
 - [14] J.F. Lawless, J. Hu, J. Cao, Methods for the estimation of failure distributions and rates from automobile warranty data, *Lifetime Data Anal.* 1 (3) (1995) 227–240.
 - [15] Y. Hong, W.Q. Meeker, Field-failure and warranty prediction based on auxiliary use-rate information, *Technometrics* 52 (2) (2010) 148–159.
 - [16] W.Q. Meeker, Y. Hong, Reliability meets big data: opportunities and challenges, *Qual. Eng.* 26 (2013) 102–116.
 - [17] M. Lu, Automotive reliability prediction based on early field failure warranty data, *Qual. Reliab. Eng. Int.* 14 (1998) 103–108.
 - [18] B. Efron, Bootstrap methods: another look at the jackknife, *Ann. Stat.* 7 (1) (1979) 1–26.
 - [19] W.Q. Meeker, L.A. Escobar, *Statistical Methods for Reliability Data*, John Wiley & Sons, Inc., New York, 1998.
 - [20] A.S. Kapadia, W. Chan, L. Moye, *Mathematical Statistics with Applications*, Taylor & Francis Group, Boca Raton, FL, 2005.
 - [21] J.A. Nelder, R. Mead, A simplex algorithm for function minimization, *Comput. J.* 7 (1965) 308–313.
 - [22] H.F. Martz, R.A. Waller, *Bayesian Reliability Analysis*, Wiley, New York, 1982.

Operational sustainability assessment of multipower source traction drive

10

I. Frenkel*, I. Bolvashenkov[†], H.-G. Herzog[†], L. Khvatkin[‡]

*Sami Shamoon College of Engineering, Beer-Sheva, Israel, [†]Technical University of Munich, Munich, Germany, [‡]Industrial Engineering and Management Department, Beer-Sheva, Israel

1 Introduction

Today with the constant growth of complexity of modern engineering systems, to achieve the required level of its sustainable and safety operations is the very complicated task. The maximum efficient and stable fulfillment of vehicles in their target functions should be considered under the operational sustainability. The implementation of the specified requirements is closely related to the assessment of the sustainable operation indicators of the system. Such of them are shown in Fig. 1.

Vehicle's traction systems are the safety-critical systems. It means that the operational sustainability of the vehicle's applications is obligatory. Bolvashenkov and Herzog [1] presented the program of the electrification of various types of vehicles, based on the electric energy, generated by renewable sources. According to this program the usage of different types of electrical traction systems in various types of vehicles now becomes very important. Particularly, correct access to the required sustainability indices for the safety-critical systems is very important.

As can be seen from Fig. 1, all indicators for sustainable operation can be divided into two groups: quantitative indicators (left column) and qualitative indicators (right column). Based on the requirements of the method of system analysis, it was necessary to take into considering the maximum possible number of local indicators, both quantitative and qualitative. Therefore a new approach is required for the solution of such complicated multipurpose problem.

Many technical systems, such as multipower source traction drive, power systems, water-cooling systems, and many others, can be described as multistate systems (MSSs). For reliability analysis of such systems the specific approaches were worked out: UGF technique [2] for steady-state performance distributions and L_Z -transform techniques [3] for dynamic MSS reliability analysis. Many technical applications of L_Z -transform are presented in Refs. [3–7].

In the present chapter, the L_Z -transform is applied to a real MS multipower source traction drive and its availability and performance are analyzed. It is shown that L_Z -transform application drastically simplifies the availability computation for such system, compared with the straightforward Markov method.

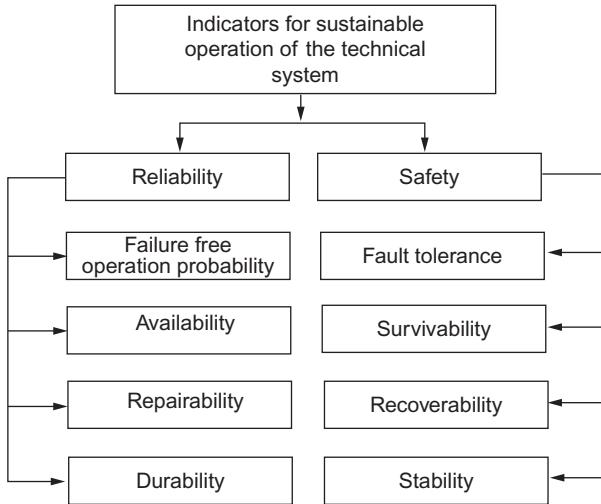


Fig. 1 Indicators for sustainable operation.

2 Brief description of the L_Z -transform method

We consider a MSS, consisting of n multistate components. Any j -component can have k_j different states, corresponding to different performances g_{ji} , represented by the set $\mathbf{g}_j = \{g_{j1}, \dots, g_{jk_j}\}$, $j = \{1, \dots, n\}$; $i = \{1, 2, \dots, k_j\}$. The performance stochastic processes $G_j(t) \in \mathbf{g}_j$ and the system structure function $G(t) = f(G_1(t), \dots, G_n(t))$ that produces the stochastic process corresponding to the output performance of the entire MSS fully define the MSS model.

The MSS model definitions can be divided into the following steps: For each multistate component we will build a model of stochastic process. Markov performance stochastic process for each component j can be represented by the expression $G_j(t) = \{\mathbf{g}_j, \mathbf{A}_j, \mathbf{p}_{j0}\}$, where \mathbf{g}_j is the set of possible component's states, defined below, $\mathbf{A}_j = (a_{lm}^{(j)}(t))$, $l, m = 1, \dots, k_j$; $j = 1, \dots, n$ —transition intensities matrix and $\mathbf{p}_{j0} = [p_{10}^{(j)} = \Pr\{G_j(0) = g_{10}\}, \dots, p_{k_j 0}^{(j)} = \Pr\{G_j(0) = g_{k_j 0}\}]$ —initial states probability distribution.

For each component j the system of Kolmogorov forward differential equations (DEs) [8] can be written for determination of the state probabilities $p_{ji}(t) = \Pr\{G_j(t) = g_{ji}\}$, $i = 1, \dots, k_j$, $j = 1, \dots, n$ under initial conditions \mathbf{p}_{j0} . Now L_Z -transform of a discrete-state continuous-time (DSCT) Markov process $G_j(t)$ for each component j can be written as follows:

$$L_Z\{G_j(t)\} = \sum_{i=1}^{k_j} p_{ji}(t) z^{g_{ji}} \quad (1)$$

The next step, in order to find L_Z -transform of the entire MSS's output performance Markov process $G(t)$, the Ushakov's universal generating operator [2] can be applied to all individual L_Z -transforms $L_Z\{G_i(t)\}$ over all time points $t \geq 0$

$$L_Z\{G(t)\} = \Omega_f\{L_Z[G_1(t)], \dots, L_Z[G_n(t)]\} = \sum_{i=1}^K p_i(t) z^{g_i} \quad (2)$$

The technique of Ushakov's operator application is well established for many different structure functions [6].

Using the resulting L_Z -transform MSS mean instantaneous availability for constant demand level, w can be derived as sum all probabilities in L_Z -transform from terms where powers of z are not negative:

$$A(t) = \sum_{g_i \geq w} p_i(t) \quad (3)$$

MSS's mean instantaneous performance may be calculated as sum of all probabilities multiplied to performance in L_Z -transform from terms where powers of z are positive:

$$E(t) = \sum_{g_i > 0} p_i(t) g_i \quad (4)$$

3 Multistate model of the multipower source traction drive

3.1 System description

We analyze a conventional diesel-electric power drive, based on a direct electric propulsion system. Such kind of power drives were used in "Amguema" type arctic cargo ships, built at the Kherson shipbuilding factory (USSR). Structure of ship's diesel-electric traction drive is shown in Fig. 2. The system consists of diesel generators subsystem, main switchboard, electric energy converter, and electric motor subsystem.

Performance of the whole system is 5500 kW. Depending on the ice conditions, the amount of cargo, and other conditions of navigation, ship's propulsion system is operated with a different number of diesel and electric propulsion motors. It carries out the required value of the performance and, as a consequence, the high survivability of the ship with the possible occurrence of critical failures of power drives equipment.

Performance of each diesel generator is 1375 kW. Therefore, connecting diesel generator in parallel supports the nominal performance, required for the functioning of whole system.

Main switchboard device and electric energy converter have nominal performance.

Performance of each electric motor is 2750 kW. To support required nominal performance for the functioning of the system, electric motors must be connecting in parallel.

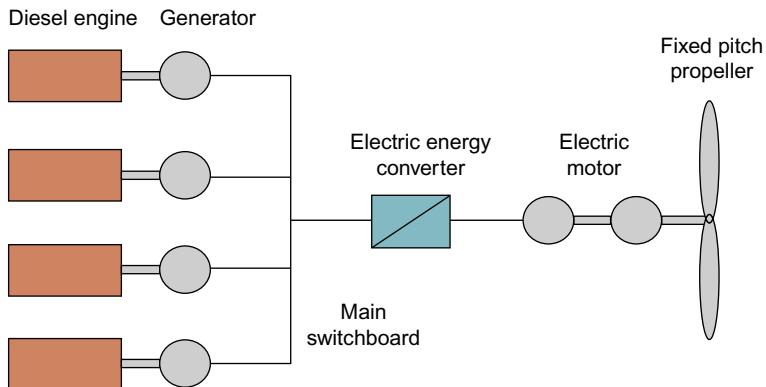


Fig. 2 Structure of the ship's diesel-electric traction drive.

In the ship's diesel-electric power drives with a fix pitch propeller, the amount of the electric machines has to be calculated accurately in order to estimate the available sufficient propulsion power, which directly determines by the required value of operational power and needed additional power in case of heavy weather or ice conditions in the area of navigation. Possible structures of the arctic ship's propulsion system with a different number of diesel generators and main traction motors are determined by operating conditions of the arctic ship and the ice and temperature conditions.

Operational modes of arctic cargo ships are as follows:

- Navigation with ice breaker in heavy ice and navigation without ice breaker in solid ice needs 100% of generated power;
- Navigation in the open water depends on required velocity, needs 75% or 50% of the generated power.

This situation is typical for the arctic cargo ships. Usually, for such systems, demand is 100%, 75%, and 50% of nominal power performance.

The state-transition diagram of the diesel-electric power drive is presented in [Fig. 3](#).

3.2 Diesel generator's subsystem

Diesel generator subsystem consists of four connected parallel pairs of diesel engines and generators. Each diesel engine and each generator are two state devices: performance of a fully operational state is 1375 kW and a total failure corresponds to a capacity of 0.

To calculate probabilities of each state we build the following system of DE for each diesel engine and each generator ($i = 1, 2, 3, 4$):

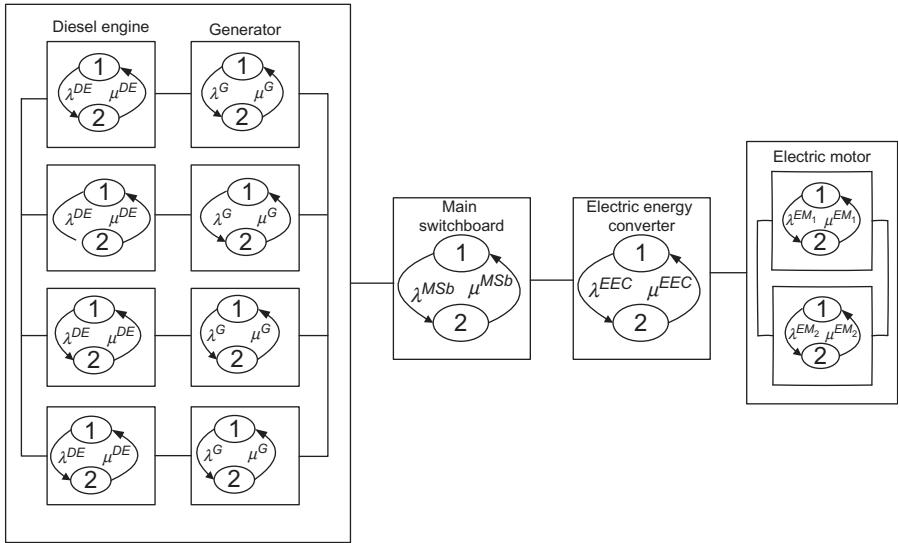


Fig. 3 State-transition diagram of the diesel-electric power drive.

$$\begin{cases} \frac{dp_{i1}^{DE}(t)}{dt} = -\lambda^{DE} p_{i1}^{DE}(t) + \mu^{DE} p_{i2}^{DE}(t), \\ \frac{dp_{i2}^{DE}(t)}{dt} = \lambda^{DE} p_{i1}^{DE}(t) - \mu^{DE} p_{i2}^{DE}(t). \end{cases} \quad \begin{cases} \frac{dp_{i1}^G(t)}{dt} = -\lambda^G p_{i1}^G(t) + \mu^G p_{i2}^G(t), \\ \frac{dp_{i2}^G(t)}{dt} = \lambda^G p_{i1}^G(t) - \mu^G p_{i2}^G(t). \end{cases}$$

Initial conditions are : $p_{i1}^{DE}(0) = 1; p_{i2}^{DE}(0) = 0$. Initial conditions are : $p_{i1}^G(0) = 1; p_{i2}^G(0) = 0$.

We used MATLAB® for numerical solution of these systems of DE to obtain probabilities $p_{i1}^{DE}(t), p_{i2}^{DE}(t)$ and $p_{i1}^G(t), p_{i2}^G(t), (i=1, 2, 3, 4)$. Therefore, for each diesel engine and generator the output performance stochastic processes can be obtained as follows:

$$\begin{cases} \mathbf{g}_i^{DE} = \{g_{i1}^{DE}, g_{i2}^{DE}\} = \{1375, 0\}, \\ \mathbf{p}_i^{DE}(t) = \{p_{i1}^{DE}(t), p_{i2}^{DE}(t)\}. \end{cases} \quad \begin{cases} \mathbf{g}_i^G = \{g_{i1}^G, g_{i2}^G\} = \{1375, 0\}, \\ \mathbf{p}_i^G(t) = \{p_{i1}^G(t), p_{i2}^G(t)\}. \end{cases}$$

Sets $\mathbf{g}_i^{DE}, \mathbf{p}_i^{DE}(t), \mathbf{g}_i^G, \mathbf{p}_i^G(t), i=1, 2, 3, 4$ define L_Z -transform for diesel engine and generator as follows:

$$\begin{aligned} L_z\{g_i^{DE}(t)\} &= p_{i1}^{DE}(t)z^{g_{i1}^{DE}} + p_{i2}^{DE}(t)z^{g_{i2}^{DE}} = p_{i1}^{DE}(t)z^{1375} + p_{i2}^{DE}(t)z^0 \\ L_z\{g_i^G(t)\} &= p_{i1}^G(t)z^{g_{i1}^G} + p_{i2}^G(t)z^{g_{i2}^G} = p_{i1}^G(t)z^{1375} + p_{i2}^G(t)z^0 \end{aligned}$$

Using the composition operator Ω_{fser} for Diesel Engine and Generator, connected in series, we obtain the L_Z -transform $L_z\{G^{DG}(t)\}$ for the diesel generator subsystem, where the powers of z are found as minimum of powers of corresponding terms:

$$\begin{aligned} L_z\{G_i^{DG}(t)\} &= \Omega_{f_{ser}}(g_i^{DE}(t), g_i^G(t)) \\ &= p_{i1}^{DE}(t)p_{i1}^G(t)z^{1375} + (p_{i1}^{DE}(t)p_{i2}^G(t) + p_{i2}^{DE}(t)p_{i1}^G(t) + p_{i2}^{DE}(t)p_{i2}^G(t))z^0. \end{aligned} \quad (5)$$

Using the following notations

$$\begin{aligned} p_{i1}^{DE-G}(t) &= p_{i1}^{DE}(t)p_{i1}^G(t); \\ p_{i2}^{DE-G}(t) &= p_{i1}^{DE}(t)p_{i2}^G(t) + p_{i2}^{DE}(t)p_{i1}^G(t) + p_{i2}^{DE}(t)p_{i2}^G(t); \end{aligned}$$

we obtain the resulting L_Z -transform for the diesel generator subsystem in the following form:

$$L_z\{G_i^{DG}(t)\} = p_{i1}^{DE-G}(t)z^{1375} + p_{i2}^{DE-G}(t)z^0, \quad i = 1, 2, 3, 4 \quad (6)$$

Using the composition operator $\Omega_{f_{par}}$ for 4 diesel generators, connected in parallel, we obtain the L_Z -transform $L_z\{G^{DG}(t)\}$ for whole diesel generator subsystem as follows:

$$\begin{aligned} L_z\{G^{DG}(t)\} &= \Omega_{f_{par}}\left(L_z\{G_1^{DG}(t)\}, L_z\{G_2^{DG}(t)\}, L_z\{G_3^{DG}(t)\}, L_z\{G_4^{DG}(t)\}\right) \\ &= p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t)z^{5500} \\ &\quad + \left[p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)\left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right) + \right] z^{4175} \\ &\quad + \left[\left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t) \right] \\ &\quad + \left[p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + \left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right) \right] z^{2750} \\ &\quad + \left[\left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right) + p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t) \right] \\ &\quad + \left[\left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t) + \right] z^{1375} \\ &\quad + \left[p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)\left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right) \right] \\ &\quad + p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)p_{32}^{DE-G}(t)p_{42}^{DE-G}(t)z^0. \end{aligned}$$

Using notations

$$\begin{aligned} P_1^{DG}(t) &= p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t); \\ P_2^{DG}(t) &= p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)\left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right) \\ &\quad + \left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t); \\ P_3^{DG}(t) &= p_{11}^{DE-G}(t)p_{21}^{DE-G}(t)p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + \left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right) \\ &\quad \cdot \left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right) + p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t); \\ P_4^{DG}(t) &= \left(p_{11}^{DE-G}(t)p_{22}^{DE-G}(t) + p_{12}^{DE-G}(t)p_{21}^{DE-G}(t)\right)p_{31}^{DE-G}(t)p_{41}^{DE-G}(t) \\ &\quad + p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)\left(p_{31}^{DE-G}(t)p_{42}^{DE-G}(t) + p_{32}^{DE-G}(t)p_{41}^{DE-G}(t)\right); \\ P_5^{DG}(t) &= p_{12}^{DE-G}(t)p_{22}^{DE-G}(t)p_{32}^{DE-G}(t)p_{42}^{DE-G}(t), \end{aligned}$$

we obtain the resulting L_z -transform for the whole diesel generator subsystem in the following form:

$$\begin{aligned} L_z\{G^{DG}(t)\} = & P_1^{DG}(t)z^{5500} + P_2^{DG}(t)z^{4125} + P_3^{DG}(t)z^{2750} \\ & + P_4^{DG}(t)z^{1375} + P_4^{DG}(t)z^0. \end{aligned} \quad (7)$$

3.3 The main switchboard

The main switchboard is a two-state device with a fully operational state with 5500 kW performance and a total failure with zero performance.

The following system of DE can be built for state probabilities calculation:

$$\begin{cases} \frac{dp_1^{MSb}(t)}{dt} = -\lambda^{MSb} p_1^{MSb}(t) + \mu^{MSb} p_2^{MSb}(t), \\ \frac{dp_2^{MSb}(t)}{dt} = \lambda^{MSb} p_1^{MSb}(t) - \mu^{MSb} p_2^{MSb}(t). \end{cases}$$

Initial conditions are: $p_1^{MSb}(0) = 1; p_2^{MSb}(0) = 0$.

A numerical solution for probabilities $p_i^{MSb}(t)$, $i = 1, 2$ can be obtained for this system of DE using MATLAB® and output performance stochastic processes for main switchboard is as follows:

$$\mathbf{g}^{MSb} \in \{g_1^{MSb}, g_2^{MSb}\} = \{5500, 0\};$$

$$\mathbf{p}^{MSb}(t) = \{p_1^{MSb}(t), p_2^{MSb}(t)\}.$$

The set $\mathbf{g}^{MSb}, \mathbf{p}^{MSb}(t)$ defines L_z -transform, associated with the main switchboard output performance stochastic process:

$$L_z\{G^{MSb}(t)\} = p_1^{MSb}(t)z^{5500} + p_2^{MSb}(t)z^0. \quad (8)$$

3.4 Electric energy converter

The electric energy converter is a two-state device with fully operational state with 5500 kW performance and a total failure corresponding to zero performance.

The system of DEs for the state for probabilities calculation is as follows:

$$\begin{cases} \frac{dp_1^{EEC}(t)}{dt} = -\lambda^{EEC} p_1^{EEC}(t) + \mu^{EEC} p_2^{EEC}(t), \\ \frac{dp_2^{EEC}(t)}{dt} = \lambda^{EEC} p_1^{EEC}(t) - \mu^{EEC} p_2^{EEC}(t). \end{cases}$$

Initial conditions are: $p_1^{EEC}(0) = 1; p_2^{EEC}(0) = 0$.

A numerical solution for probabilities $p_i^{EEC}(t)$, $i=1,2$ can be obtained using MATLAB®. The electric energy converter output performance stochastic process is as following:

$$\begin{aligned}\mathbf{g}^{EEC} &\in \{g_1^{EEC}, g_2^{EEC}\} = \{5500, 0\}; \\ \mathbf{p}^{EEC}(t) &= \{p_1^{EEC}(t), p_2^{EEC}(t)\}.\end{aligned}$$

The set \mathbf{g}^{EEC} , $\mathbf{p}^{EEC}(t)$ defines L_z -transform, associated with the electric energy converter output performance stochastic process:

$$L_z\{G^{EEC}(t)\} = p_1^{EEC}(t)z^{5500} + p_2^{EEC}(t)z^0. \quad (9)$$

3.5 Electric motors subsystem

As it was presented above, performance of each electric motor is 2750 kW, half of the capacity of the system. Therefore, to receive the nominal performance, which is required for the functioning of the system, electric motors have to be connected in parallel. Each electric motor can be in one of two states: an operational state with a performance of 2750 kW and a total failure state corresponding to zero performance.

Using the Markov method and according to state-transitions diagram (Fig. 3), the following system of DE for each electric motor ($i=1, 2$) can be built:

$$\begin{cases} \frac{dp_{i1}^{EM}(t)}{dt} = -\lambda^{EM} p_{i1}^{EM}(t) + \mu^{EM} p_{i2}^{EM}(t), \\ \frac{dp_{i2}^{EM}(t)}{dt} = \lambda^{EM} p_{i1}^{EM}(t) - \mu^{EM} p_{i2}^{EM}(t). \end{cases}$$

Initial conditions are: $p_{i1}^{EM}(0) = 1$; $p_{i2}^{EM}(0) = 0$.

A numerical solution for probabilities $p_{i1}^{EM}(t)$ and $p_{i2}^{EM}(t)$ ($i=1, 2$) can be obtained for each of these two systems of DE using MATLAB®. Therefore, for each electric motor we obtain the output performance stochastic processes:

$$\begin{cases} \mathbf{g}_i^{EM} = \{g_{i1}^{EM}, g_{i2}^{EM}\} = \{2750, 0\}, \\ \mathbf{p}_i^{EM}(t) = \{p_{i1}^{EM}(t), p_{i2}^{EM}(t)\}. \end{cases}$$

The sets \mathbf{g}_i^{EM} , $\mathbf{p}_i^{EM}(t)$, $i=1,2$ define for each electric motor the following L_z -transform:

$$\begin{aligned}L_z\{g_i^{EM}(t)\} &= p_{i1}^{EM}(t)z^{g_{i1}^{EM}} + p_{i2}^{EM}(t)z^{g_{i2}^{EM}} \\ &= p_{i1}^{EM}(t)z^{2750} + p_{i2}^{EM}(t)z^0.\end{aligned} \quad (10)$$

Using the composition operator $\Omega_{f_{par}}$ for two electric motors, connected in parallel, the L_Z -transform $L_z\{G^{EM}(t)\}$ for the electric motors subsystem can be obtained in the following form:

$$\begin{aligned} L_z\{G^{EM}(t)\} &= \Omega_{f_{par}}(L_z\{g_1^{EM}(t)\}, L_z\{g_2^{EM}(t)\}) \\ &= \Omega_{f_{par}}(p_{11}^{EM}(t)z^{2750} + p_{12}^{EM}(t)z^0, p_{21}^{EM}(t)z^{2750} + p_{22}^{EM}(t)z^0) \\ &= p_{11}^{EM}(t)p_{21}^{EM}(t)z^{5500} + (p_{11}^{EM}(t)p_{22}^{EM}(t) + p_{21}^{EM}(t)p_{12}^{EM}(t))z^{2750} + p_{12}^{EM}(t)p_{22}^{EM}(t)z^0. \end{aligned}$$

Using the following notations,

$$\begin{aligned} P_1^{EM}(t) &= p_{11}^{EM}(t)p_{21}^{EM}(t); \\ P_2^{EM}(t) &= p_{11}^{EM}(t)p_{22}^{EM}(t) + p_{21}^{EM}(t)p_{12}^{EM}(t); \\ P_3^{EM}(t) &= p_{12}^{EM}(t)p_{22}^{EM}(t); \end{aligned}$$

L_Z -transform for the electric motors subsystem can be presented in the following form:

$$L_z\{G^{EM}(t)\} = P_1^{EM}(t)z^{5500} + P_2^{EM}(t)z^{2750} + P_3^{EM}(t)z^0. \quad (11)$$

3.6 Multistate model for multipower source traction drive

As it was presented in Fig. 3, all sub-systems and elements of the system are connected in series. Therefore, the whole system L_Z -transform is as follows:

$$\begin{aligned} L_z\{G^{MPD}(t)\} &= \Omega_{f_{ser}}(L_z\{G^{DG}(t)\}, L_z\{G^{MSb}(t)\}, L_z\{G^{EBC}(t)\}, L_z\{G^{EM}(t)\}) \\ &= \Omega_{f_{ser}}(P_1^{DG}(t)z^{5500} + P_2^{DG}(t)z^{4125} + P_3^{DG}(t)z^{2750} + P_4^{DG}(t)z^{1375} + P_4^{DG}(t)z^0, \\ &\quad p_1^{MSb}(t)z^{5500} + p_2^{MSb}(t)z^0, p_1^{EBC}(t)z^{5500} + p_2^{EBC}(t)z^0, P_1^{EM}(t)z^{5500} + P_2^{EM}(t)z^{2750} + P_3^{EM}(t)z^0) \end{aligned}$$

Using simple algebra calculations of the powers of z as minimum values of powers of corresponding terms, the whole system's L_Z -transform expression is as following:

$$\begin{aligned} L_z\{G^{MPD}(t)\} &= P_1^{MPD}(t)z^{5500} + P_2^{MPD}(t)z^{4125} + P_3^{MPD}(t)z^{2750} \\ &\quad + P_4^{MPD}(t)z^{2250} + P_5^{MPD}(t)z^{1375} + P_6^{MPD}(t)z^0, \end{aligned} \quad (12)$$

where

$$\begin{cases} g_1^{MPD} = 5500 \text{kW}, \\ P_1^{MPD}(t) = P_1^{DG}(t)p_1^{MSb}(t)p_1^{EBC}(t)P_1^{EM}(t), \end{cases}$$

$$\begin{cases} g_2^{MPD} = 4125 \text{kW}, \\ P_2^{MPD}(t) = P_2^{DG}(t)p_1^{MSb}(t)p_1^{EBC}(t)P_1^{EM}(t); \end{cases}$$

$$\begin{cases} g_3^{MPD} = 2750 \text{kW}, \\ P_3^{MPD}(t) = P_3^{DG}(t)p_1^{MSb}(t)p_1^{EBC}(t)P_1^{EM}(t); \end{cases}$$

$$\begin{cases} g_4^{MPD} = 2250 \text{kW}, \\ P_4^{MPD}(t) = (P_1^{DG}(t) + P_2^{DG}(t) + P_3^{DG}(t))p_1^{MSb}(t)p_1^{EBC}(t)P_2^{EM}(t); \end{cases}$$

$$\begin{cases} g_5^{MPD} = 1375 \text{kW}, \\ P_5^{MPD}(t) = P_4^{DG}(t)p_1^{MSb}(t)p_1^{EBC}(t); \end{cases}$$

$$\begin{cases} g_6^{MPD} = 0 \text{kW}, \\ P_5^{MPD}(t) = (P_1^{DG}(t) + P_2^{DG}(t) + P_3^{DG}(t) + P_4^{DG}(t))p_1^{MSb}(t)(p_1^{EBC}(t)P_3^{EM}(t) + p_1^{EBC}(t)) \\ + p_2^{MSb}(t) + P_5^{DG}(t)p_1^{MSb}(t); \end{cases}$$

Using expression (12), the MSS mean instantaneous availability for constant demand level w may be presented as follows:

- For 100% demand level ($w=5500 \text{ kW}$)

$$A_{w \geq 5500 \text{ kW}}(t) = \sum_{g_k \geq 5500} P_k^{MPD}(t) = P_1^{MPD}(t) \quad (13)$$

- For 75% demand level ($w=4125 \text{ kW}$)

$$A_{w \geq 4125 \text{ kW}}(t) = \sum_{g_k \geq 4125} P_k^{MPD}(t) = \sum_{k=1}^2 P_k^{MPD}(t) = P_1^{MPD}(t) + P_2^{MPD}(t); \quad (14)$$

- For 50% demand level ($w=2750 \text{ kW}$)

$$A_{w \geq 2750 \text{ kW}}(t) = \sum_{g_k \geq 2750} P_k^{MPD}(t) = \sum_{k=1}^3 P_k^{MPD}(t) = P_1^{MPD}(t) + P_2^{MPD}(t) + P_3^{MPD}(t) \quad (15)$$

The MSS instantaneous mean power performance for the multipower source traction drive can be obtained as follows:

$$E(t) = \sum_{g_i > 0} P_i^{MPD}(t)g_i^{MPD} = \sum_{i=1}^5 P_i^{MPD}(t)g_i^{MPD}. \quad (16)$$

4 Power performance calculation

The failure and repair rates (in year⁻¹) of each system's elements are presented in **Table 1**.

MSS mean instantaneous availability for different constant demand levels is presented in **Fig. 4**.

As one can see from **Fig. 4**, the instantaneous availability depends on operation conditions and changes from 88% for 100% of the required generated power to 98.4% for 50% of required generated power.

Calculated MSS instantaneous mean power performance of the multipower source traction drive is presented in **Fig. 5**.

Table 1 Failure and repair rates of each system's elements

	Failure rates	Repair rates
Diesel engine	4.99	180
Generator	0.006	50
Main switchboard	0.006	20
Electric energy converter	0.44	25
Electric motor	0.0026	100

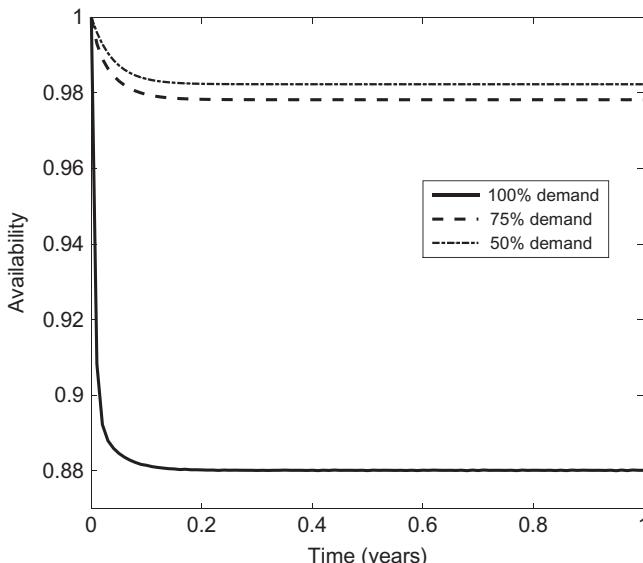


Fig. 4 MSS mean instantaneous availability for different constant demand levels.

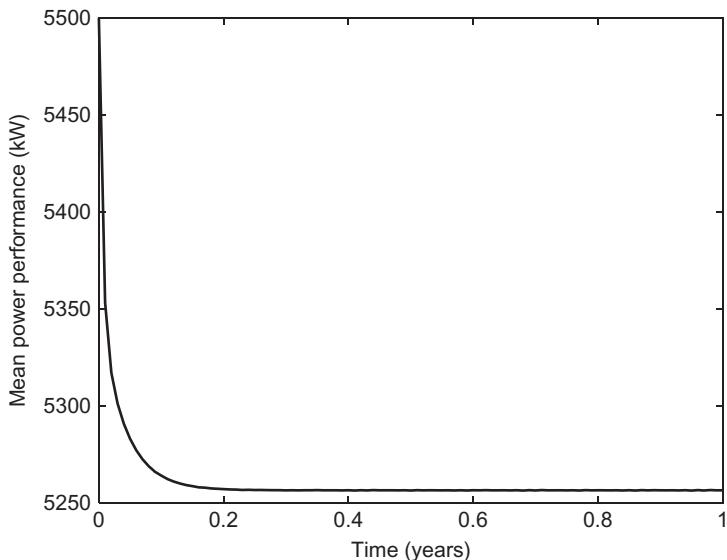


Fig. 5 MSS instantaneous mean power performance of the multipower source traction drive.

5 Conclusion

In this chapter, the L_Z -transform method was used for evaluation of two important parameters of the vehicle's operational sustainability—availability and performance of the multistate multipower source traction drive.

L_Z -transform approach extremely simplifies the solution, which in comparison with straightforward Markov method would have required building and solving the model with 4096 states.

The proposed approach allows an optimization of the number of the power sources of traction drive, their characteristics, and schemes of connection in terms of providing the maximum operational sustainability.

References

- [1] I. Bolvashenkov, H.-G. Herzog, Use of stochastic models for operational efficiency analysis of multi power source traction drives. in: I. Frenkel, A. Lisnianski (Eds.), Proceedings of the Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, SMRLO'16, Beer Sheva, Israel, February 15–18, IEEE CPS, 2016, pp. 124–130, <http://dx.doi.org/10.1109/SMRLO.2016.30>.
- [2] I. Ushakov, A universal generating function, Sov. J. Comput. Syst. Sci. 24 (1986) 37–49.
- [3] A. Lisnianski, L_Z -transform for a discrete-state continuous-time Markov process and its application to multi-state system reliability, in: A. Lisnianski, I. Frenkel (Eds.), Recent

- Advances in System Reliability. Signatures Multi-State Systems and Statistical Inference, Springer, London, 2012, pp. 79–95.
- [4] I. Frenkel, I. Bolvashenkov, H.-G. Herzog, L. Khvatskin, Performance availability assessment of combined multi power source traction drive considering real operational conditions. *Transport Telecommun.* 17 (3) (2016) 179–191, <http://dx.doi.org/10.1515/ttj-2016-0016>.
 - [5] P. Lein, I. Frenkel, L. Khvatskin, On availability determination for mss cold water supply system by Lz-transform: case study. in: I. Frenkel, A. Lisnianski (Eds.), Proceedings of the Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, SMRLO'16, Beer Sheva, Israel, February 15–18, IEEE CPS, 2016, pp. 365–374, <http://dx.doi.org/10.1109/SMRLO.2016.65>.
 - [6] A. Lisnianski, I. Frenkel, Y. Ding, Multi-State System Reliability Analysis and Optimization for Engineers and Industrial Managers, Springer, London, 2010.
 - [7] B. Natvig, Multistate Systems Reliability. Theory With Applications, Wiley, New York, 2011.
 - [8] K. Trivedi, Probability and Statistics with Reliability, Queuing and Computer Science Applications, Wiley, New York, 2002.

Index

Note: Page numbers followed by *f* indicate figures, and *t* indicate tables.

A

- Absorber type muffler, 77–78
- Acoustic impedance, 65–67
- Acoustics
 - application of, 80–81
 - definition, 55–56
 - imaging, 81
 - lens, 80
 - Linear Euler equation, 56–57
 - sound propagation (*see* Sound wave propagation)
 - wave equation, 57–59
- AISI 4130 alloy, 136
 - FEA results, 138
 - and gray cast iron materials, 136, 137*f*, 138
 - mechanical properties, 135
 - meshed model of, 134–135
 - mode shape and natural frequency of, 137*f*
 - vibration mode shapes, 140–142
- Anderson-Darling (AD) statistics, 174, 179–180
- Appraisal reliability, of sliding window system, 85–88
- Availability
 - CTSMMs, 29–32, 43*t*, 45*f*
 - DTSMMs, 32–34, 46–47*t*, 48*f*
 - series-parallel system, 102–103, 103*t*, 104*f*, 111
- Axial bending vibration, 136, 142–143

B

- Band gap, 75–77
- Barlow-Proschan index, 84, 88, 93–94
- Bayesian analysis, 183, 186–188
- Bloch–Floquet theorem, 77
- Bootstrap resampling algorithm, 175–176, 179–180

C

- CAD model, 134–135, 134*f*
- Cavity, sound propagation in, 71–73, 71*f*
- Chapman–Kolmogorov equation, 7
- CIDTS. *See* Cooperative intrusion detection and tolerance system (CIDTS)
- Commercial-off-the-shelf (COTS), 22–26, 46–49
- Compression of wave, 55–56
- Conical horn, sound propagation, 61–62, 61*f*
- Connecting bolt (CB) constraint analysis, 138–144, 139*t*, 139–145*f*
- Consecutive-*k*-out-of-*n* system, 1–2
- Consecutive-*k* system, 2–4, 7
- Consecutive success total failure (CSTF) test models, 3
- Continuous-time semi-Markov models (CTSMMs), 22–23
 - MTTSF, 44*t*, 45*f*
 - optimal PPMT, 29–32
 - security evaluation with, 42–45
 - steady-state system availability, 29–32, 43*t*, 45*f*
 - transition diagram of, 27, 27*f*
- Cooperative intrusion detection and tolerance system (CIDTS), 21–22
- COTS. *See* Commercial-off-the-shelf (COTS)
- Crowding distance technique, 119*f*.
See also Multi-objective particle swarm optimization incorporating crowding distance (MOPSO-CD)
- CSTF test models. *See* Consecutive success total failure (CSTF) test models
- CTSMMs. *See* Continuous-time semi-Markov models (CTSMMs)

MATHEMATICS APPLIED TO ENGINEERING

EDITED BY

MANGEY RAM AND J. PAULO DAVIM

A comprehensive guide to the different mathematical techniques used in many fields of engineering, such as robotics, computer science and mechanical engineering.

- Covers many mathematical techniques for robotics, computer science, mechanical engineering, HCl, and Machinability.
- Describes different algorithms.
- Explains different modelling techniques and simulations.

Mathematics Applied in Engineering presents a wide array of applied mathematical techniques for an equally wide range of engineering applications, covering areas such as acoustics, system engineering, optimization, mechanical engineering, and reliability engineering.

Mathematics acts as a foundation for new advances, as engineering evolves and develops. This book will be of great interest to postgraduate and senior undergraduate students, and researchers, in engineering and mathematics, as well as to engineers, policy makers, and scientists involved in the application of mathematics in engineering.

About the Editors

Mangey Ram is currently Head of the Department and Assistant Dean (International Affairs) at Graphic Era University, India. He is Editor-in-Chief of the International Journal of Mathematical, Engineering and Management Sciences; Executive Editor, Journal of Graphic Era University; Executive Associate, Journal of Reliability and Statistical Studies. Prof. Ram has published 100 plus research publications and presented his works at national and international levels. His research interests are in reliability theory and applied mathematics.

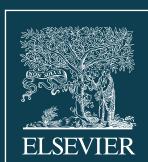
J. Paulo Davim is Professor at the Department of Mechanical Engineering, University of Aveiro, Portugal. He has more than 30 years of teaching and research experience in manufacturing, materials and mechanical engineering. He is the Editor-in-Chief of several international journals, and is currently an editorial board member of 30 international journals. In addition, he has published in his field of research, as author and co-author, more than 60 book chapters and 400 articles for journals and conferences.

Related Titles

Kundu & Ganguly | Analysis of Step-Stress Models: Existing Results and Some Recent Developments | Academic Press | 9780128097137

Yang | Engineering Mathematics with Examples and Applications | Academic Press | 9780128097304

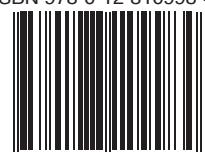
Blinder | Guide to Essential Math: A Review for Physics, Chemistry and Engineering Students, Second Edition | Elsevier | 9780124071636



ACADEMIC PRESS

An imprint of Elsevier
elsevier.com/books-and-journals

ISBN 978-0-12-810998-4



9 780128 109984

MATHEMATICS APPLIED TO ENGINEERING

EDITED BY
MANGEY RAM AND J. PAULO DAVIM

A comprehensive guide to the different mathematical techniques used in many fields of engineering, such as robotics, computer science and mechanical engineering.

- Covers many mathematical techniques for robotics, computer science, mechanical engineering, HCI, and Machinability.
- Describes different algorithms.
- Explains different modelling techniques and simulations.

Mathematics Applied in Engineering presents a wide array of applied mathematical techniques for an equally wide range of engineering applications, covering areas such as acoustics, system engineering, optimization, mechanical engineering, and reliability engineering.

Mathematics acts as a foundation for new advances, as engineering evolves and develops. This book will be of great interest to postgraduate and senior undergraduate students, and researchers, in engineering and mathematics, as well as to engineers, policy makers, and scientists involved in the application of mathematics in engineering.

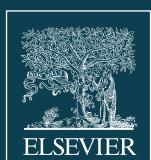
About the Editors

Mangey Ram is currently Head of the Department and Assistant Dean (International Affairs) at Graphic Era University, India. He is Editor-in-Chief of the International Journal of Mathematical, Engineering and Management Sciences; Executive Editor, Journal of Graphic Era University; Executive Associate, Journal of Reliability and Statistical Studies. Prof. Ram has published 100 plus research publications and presented his works at national and international levels. His research interests are in reliability theory and applied mathematics.

J. Paulo Davim is Professor at the Department of Mechanical Engineering, University of Aveiro, Portugal. He has more than 30 years of teaching and research experience in manufacturing, materials and mechanical engineering. He is the Editor-in-Chief of several international journals, and is currently an editorial board member of 30 international journals. In addition, he has published in his field of research, as author and co-author, more than 60 book chapters and 400 articles for journals and conferences.

Related Titles

- Kundu & Ganguly | Analysis of Step-Stress Models: Existing Results and Some Recent Developments | Academic Press | 9780128097137
- Yang | Engineering Mathematics with Examples and Applications | Academic Press | 9780128097304
- Blinder | Guide to Essential Math: A Review for Physics, Chemistry and Engineering Students, Second Edition | Elsevier | 9780124071636



ACADEMIC PRESS
An imprint of Elsevier
elsevier.com/books-and-journals



MATHEMATICS APPLIED TO ENGINEERING

RAM • DAVIM

ACADEMIC PRESS

MATHEMATICS APPLIED TO ENGINEERING

EDITED BY
**MANGEY RAM
J. PAULO DAVIM**

